Language Competence and Working Memory in Older Adults

A thesis
submitted in partial fulfillment
of the requirements for the Degree of
Masters of Science in Psychology
in the
University of Canterbury
by

Elena Loukavenko

University of Canterbury
2002

CONTENTS

CONTENTS (continued)

PAGE

CONTENTS (continued)

CONTENTS (continued)

PAGE

LIST OF FIGURES AND TABLES

FIGURE **PAGE**

LIST OF FIGURES AND TABLES (continued)

TABLE **PAGE**

LIST OF FIGURES AND TABLES (continued)

PAGE

TABLE

LIST OF FIGURES AND TABLES (continued)

PAGE

TABLE

LIST OF FIGURES AND TABLES (continued)

PAGE

TABLE

ACKNOWLEDGEMENTS

I would like to dedicate this work to my mother Dr. Tatiana Blagova who has always encouraged my interests in the field of neuropsychology.

I would like to express my appreciation of the time spent with me, and encouragement provided to me by my supervisor Dr John Dalrymple-Alford. His expertise, advice and guidance were invaluable and very much appreciated throughout this project.

Very grateful thanks also to Prof. Garth Fletcher, Dr Bruce Ellis and Dr Paul Barrett for their assistance with statistical analysis.

I also want to thank Catherine Moran for sharing with me her knowledge of linguistics and psychometric test administration. My thanks also to Dr Meredyth Daneman who so kindly provided the items and instructions for the reading span task administration and to Dr Brenda Hannon for her valuable comments on testing procedures.

Much thanks to Frederic Durel for promptly designing and implementing an excellent computer program for data recording, and also to John Barton for technical support.

My heart felt thanks to my husband James Moran for his support and understanding.

Last but not least my very grateful thanks to all the people who so willingly donated their time to take part in this research. I would like to thank the members of the Pegasus Lions Club, in particular Shirlee Hamilton and Julie Marsh for their assistance with participant recruitment.

ABSTRACT

Performance on the Test of Language Competence-Expanded Edition (TLC-E), which measures discourse and higher language skills, was assessed in community-based representative groups of young adults (20-34 years; reference group), young-old adults (65-74 years) and old-old adults (75-89 years). Clear evidence of progressive age-related deterioration wasobtained on the TLC-E composite score, as well as the Ambiguous Sentences, Making Inferences and Recreating Sentences subtests. The Figurative Language subtest, based on metaphoric expressions, revealed a decline in the old-old elderly only. Given these differences and the clinical utility of the TLC-E as a tool for assessment of linguistic impairment, this study provides provisional norms for this test (i.e. for each of the young, young-old and old-old groups). Age-related declines were also identified in terms of several measures of working memory, processing speed, Stroop inhibition and long-term memory, which were examined to determine their contribution to the effects of aging on language performance. Path analyses indicated that the contribution of speed and long-term memory to differences on the TLC-E composite score was only indirect and was mediated by working memory, which itself also directly explained age-related differences on this language measure. Stroop Inhibition was not associated with language performance. Individual TLC-E subtests, however, were differentially associated with the cognitive predictor variables examined. These findings provide new evidence on impaired discourse skills in the elderly and extend our understanding of the relationship between aging and cognition by clarifying the important role of working memory by comparison with other cognitive processes as potential mediators of changes in higher language skills in the elderly.

1. INTRODUCTION

1.1 Brief Outline of the Present Research

The primary objective of the present research was to compare the performance of younger and older adults on the Test of Language Competence – Expanded Edition (TLC-E; Wiig & Secord, 1989), which is principally a clinical test of discourse language and higher language skills. The purpose of this comparison was to establish whether healthy elderly individuals exhibit any decline on these tasks of complex language functioning. If any decline was the case the obtained results would also provide a major contribution to the establishment of provisional norms for the TLC-E for the healthy New Zealand elderly population in two age bands (65-74 and 75-89 years) and in a young "reference group" (20-34 years). By dividing the elderly sample into groups of young-old (65-74 years) and old-old (75-89 years) the question of possible changes in language and other cognitive abilities with advancing age was also addressed. It was envisaged that the obtained data would be of considerable value for both future research and clinical purposes, in terms of providing comparative information for the evaluation of performance of individuals with neurodegenerative disorders as well as the healthy elderly. Such comparisons have high clinical utility as they serve to enhance our understanding of individual differences among the elderly and to differentiate the normal age-related decline from the start of any dementia process.

The second main objective of the present research was to identify other variables that may contribute to performance on the TLC-E. The participants were tested on a selected range of working memory, processing speed, inhibitory efficiency and long-term verbal memory tasks, on which the elderly have often been reported to show mild to moderate deficits. The speed with which the operations are conducted is thought to contribute to the age-related differences observed on measures of cognitive functioning, including language. While inhibitory efficiency may affect language functioning by allowing the individual to ward off distractions and inhibit irrelevant thoughts. Long-term memory abilities that reflect the capacity to learn and retain information may also be a factor which affects language functioning. Most importantly, the research focused on evaluating whether any decline in language competence in the elderly is primarily influenced by age-related changes in working memory.

1.2 Test of Language Competence-Expanded Edition as a Measure of Higher Language Skills

Limited normative data exist regarding which aspects of higher language functioning are preserved or impaired in the elderly (Snowdown, 1997). Recent reviews (Caplan & Waters, 1999, Burke & Mackay, 1997, Kemper & Kliegl, 1999, Wingfield & Stine-Morrow, 2000) indicate that higher language function, including discourse skills and the pragmatic use of language, is more likely than non-discourse language to show differences in older compared to younger individuals. In particular older adults were found to experience difficulties in both producing (Kemper, 1987) and comprehending (Burke & Harold, 1988) syntactically complex sentences, ascertaining the meaning of metaphoric expressions (Light, Owens, Mahoney & La Voie, 1993), comprehending ambiguous statements (Kellas, Paul, & Vu, 1995) and drawing inferences from text (Hasher & Zacks, 1988). These age differences on discourse tasks tend to become more profound when the tasks at hand require the participants to memorize large amounts of information, to manipulate this information in memory or to perform the task under time restraint (Kemper & Mitzner, 2001). In contrast, when discourse skills are examined using stimuli which consist of meaningful language, within its natural linguistic context and with an opportunity to review the information, age differences become less apparent (Caplan & Waters, 1999). There is currently a major need for normative information on performance of elderly on higher language tasks that approximate the natural discourse situations and do not place a heavy demand on memory or speed of responding. The Test of Language Competence-Expanded Edition (Wiig & Secord, 1989) possesses a number of characteristics that make it optimally suitable for evaluation of discourse skills in the elderly.

The Test of Language Competence –Expanded Edition (TLC-E) (Wiig & Secord, 1989) was designed to examine the discourse and higher language in children aged 9 and above and adults with higher language disabilities. It assesses language competence, in terms of appropriate understanding and/or expression of language content and response to the communication demands of specific situations. Spontaneous natural interaction between language users does not take place unless language content and function (communicative intent) are integrated in a specific communicative situation (context). The TLC-E was designed to measure language content in a communicative context. The four TLC-E subtests were constructed to sample a variety of expressions (language content), some highly literal and others highly figurative, across a series of contexts (communicative situations), some fairly restrictive and some quite interactive. The TLC-E subtests assess the use of language at the level of semantics, semantic-syntactic interfaces and pragmatics, and test the use of propositions in narrow or communication-like contexts. The subtests of the TLC-E also demand problem solving, planning, and decision making with alternative solutions or responses to the same linguistic input.

The first of the four subtests of the TLC-E, *Ambiguous Sentences*, requires multiple interpretations of sentences and reflects propositions in narrow contexts. Ambiguous Sentences would generally go unnoticed in everyday conversations that occur within a context and feature a topical focus, as the topic and context determine the interpretation. When reading, however, ambiguous sentences may require greater effort of interpretation, as the context or topical reference may not be obvious to the reader and may be conceptually removed. Hence, when ambiguous sentences are encountered the reader first has to compute all the possible meanings. After all the possible meanings are derived and pragmatic/syntactic biases are considered the reader makes a selection of a single meaning. A tendency to concentrate on a single meaning (probably highly concrete or experience based) before realizing the possibility of an alternative may result in a failure to resolve the ambiguity.

The second subtest, *Making Inferences*, requires the person to make plausible inferences on the basis of two sentences that describe the lead-in and lead-out of a causal event chain. The Making Inferences subtest reflects propositions in narrow context. The ability to make inferences or understand implied but not stated relationships, relies on recognition and recovery of missing links in the underlying causal chain of a script. The scripts are stored in episodic memory and are further organized by scenes that comprise what is general about the scripts. The failure of the listener/reader to recognize and retrieve the scripts will affect their ability to resolve the inference.

The third subtest, *Recreating Sentences*, requires reconstruction of propositions in communication-like context, and asks the person to recreate a sentence that could have been said by one of the people in a presented illustration, incorporating three given lexical units. In order to accomplish the task the person has to 1) interpret each of the stimulus words/concepts, 2) discern the grammatical function/case of each concept, 3) formulate a grammatically correct sentence incorporating the concepts and 4) relate the sentence to the situational context. This task requires high-level integration of variables related to content, form and use.

The last subtest, *Figurative Language*, requires the interpretation of figurative language and metaphoric expressions and reflects the use of propositions in a communication-like context. When sufficient prior context is available to the listener/reader, the intended non-literal meaning is thought to become clear through the activation of relevant schemata. On the contrary, when no sufficient context is presented the individual is forced firstly to compute the literal meaning fully before comprehending the non-literal. The ability of the listener/reader to interpret metaphors is also related to the familiarity of the metaphor as well as the degree of connection between the literal and the non-literal meaning of the metaphor.

This summary of the TLC-E shows that it assesses discourse skills in a format that closely approximates natural discourse constraints, which sets it apart from most other un-natural, overly complex language tasks that may not accurately reflect language capabilities under normal discourse conditions. Several further other advantages also make the TLC-E a valuable tool in language assessment. The TLC-E utilizes a testing procedure that minimizes the amount of information the participant is required to remember by presenting all the stimulus material in written as well as verbal form. Hence, it provides the person with ample opportunity to review the material and de-emphasizes speed of responding, providing an evaluation of language functioning without the interference of high memory load or high speed demands. Furthermore, by presenting the information to the participant in both written and oral form the TLC-E minimizes the possible negative affects presbycusis (age related loss of hearing acuity) might have on comprehension.

The TLC-E also has high clinical utility, which is not characteristic of most language tests used in research. In recent years, some researchers have begun to use the TLC-E to examine the higher language abilities of patients suffering from neurodegenerative disorders. Research with multiple sclerosis sufferers indicated that these patients experienced significantly more difficulties than the control group on all of the TLC-E subtests (Lethlean & Murdoch, 1997). The same results were obtained with Alzheimer's dementia patients (Harris, 1994) who were found to be particularly impaired on the Making Inferences and Recreating Sentences subtests. A study with Parkinson's disease sufferers indicated that these patients have difficulties interpreting ambiguous and figurative language (Lewis, Lapointe, Murdoch & Chenery, 1998). The interpretation of findings from these studies would have been greatly assisted if normative information for performance of adults and elderly on the TLC-E existed. The studies either used comparisons based on children's norms provided by the TLC-E (Harris, 1994) or relied on a comparison with a control group. Control group comparisons are obviously valuable, but may be misleading due to often unrepresentative nature of the sample. As shown in Table 1, these controls often generate variable results relative to the manual's standards, which may reflect the composition of the age group used. Hence an investigation that aims to collect data on performance of a representative sample of healthy elderly on the TLC-E is well overdue.

In summary, the TLC-E's ability to test discourse language in communication-like context, its limited demand on memory or speed of processing, and its high clinical utility make it an optimal test for assessment and comparison of language skills in the elderly.

Table 1. Raw score means and standard deviations obtained for control groups in the Lewis, et al (1998) and Leathlean & Murdoch (1997) studies, and those provided by the TLC-E Manual.

	Wiig & Secord (1989) TLC-E manual	Lewis et. al (1998)	Lethlean & Murdoch (1997)
	(17-18+ years) (n=112)	(51-85 years) (n=60)	(26-76 years) (n=20)
	M SD	M SD	M SD
Task			
TLC-E	167.3 (13.8)	*	174.9 (7.3)
AS	32.6 (5.1)	34.0 (3.0)	35.3 (3.0)
MI	32.8 (3.0)	*	31.1 (3.1)
RS	71.2 (6.0)	73.2 (5.2)	75.6 (2.6)
FL	30.7 (4.9)	31.5 (3.4)	33.2 (2.3)

Note: TLC-E= TLC- E composite score, AS- Ambiguous sentences, MI= Making Inferences, RS= Recreating sentences, FL= Figurative Language, *= no score provided.

1.3 Language Functioning in the Elderly

Cognitive aging researchers have documented a number of age-related declines in older adults' language processing. Age-related differences have been observed in auditory-verbal discourse comprehension (Ulatowska, Cannito, Hayashi & Fleming, 1986), written language comprehension (Light & Anderson, 1985), capacity for recall of written language (Light, 1990), reconstruction (Ulatowska, Cannito, Hayashi & Fleming, 1986) and inferencing (Kemper & Anagnopoulos, 1993). As it is beyond the scope of this thesis to discuss all the findings in the field, this introduction will concentrate on what is currently known about how normal aging affects performance on tasks similar to those assessed by the TLC-E.

Discourse encompasses a variety of communication skills ranging from opening and closing conversations, maintaining and shifting topics, telling stories and even modifying personal relations. At the core of effective discourse production lies an ability to construct a grammatically and semantically correct and coherent sentence. This ability is well captured by the Recreating Sentences subtest of the TLC-E. The study of language production in the elderly has been relatively neglected in the literature. The majority of research in the field was conducted by Kemper and colleagues who argue that older people tend to produce syntactically less complex sentences and exhibit higher number of inaccuracies in discourse (Kemper, 1992). For example, examination of diary entries of older adults over a period of 10 years indicated a decline in syntactic complexity in terms of the number of clauses, types of subordinates, gerunds, and double and triple embeddings used (Kemper, 1987).

By the age of 70 years people tended to use less complicated sentences, averaging about one clause, whereas in their 20's their sentences averaged about three clauses. However, the length of the sentence did not alter with years, supporting the notion that older adults tend to produce less syntactically complex sentences. In the same vein, Bromley (1991) also found elderly adults to write less complex sentences than younger adults when writing self-descriptions (but there was no age difference in either sentence length or readability). Similar findings are obtained when speech production is investigated. Kynette and Kemper (1986) studied the spontaneous speech of 50-90 year olds and found that people in their seventies and eighties were less likely to use left-branching constructions (a construction where an embedded clause occurs to the left of the main clause for e.g. *"The girl who runs the nursery school for our church is awfully young*".) and that many of their grammatical errors occurred when attempting to use such constructions. Furthermore, older adults had significant difficulties imitating syntactically complex left-branching sentences, although they had minimal difficulties producing and imitating less complex right-branching sentences. Nevertheless, Kemper affirmed that: "Although studies of adults' speech production and writing revealed that older adults are unlikely to produce complex grammatical forms spontaneously, their speech does not evidence progressive degeneration into 'baby talk'" (1992, p.221). This relative maintenance of language ability has been attributed to the ability of the elderly to effectively utilize the discourse context in order to overcome production difficulties (Kemper and Anagnopoulos, 1993).

Whilst age related changes are usually present when discourse production is examined the same is not true for discourse comprehension, which presents a more complex picture of sparing and impairment. Assessment of one's ability to interpret ambiguous sentences is commonly used when age-related changes in comprehension are examined. Two main experimental paradigms are normally utilized in this case. One approach involves evaluating comprehension on-line (as it occurs) and the other involves assessing comprehension off-line by asking comprehension questions. When ambiguity comprehension is examined on-line age-related declines are not always observed. For example, examination of ambiguity identification in younger and older adults as single words and sentences has not been found to produce age differences in the ability of participants to identify single ambiguous words on-line (Kellas, Paul, and Vu, 1995). Kemtes & Kemper (1997) examined younger and older adults' on-line comprehension of temporarily ambiguous sentences *(Several angry workers warned about low wages…)* that was resolved either with the main verb interpretation (*during the holiday season*) or a relative clause interpretation (*decided to file a complaint*). The researchers also assessed adults' off-line comprehension by presenting comprehension questions after each sentence was read. The main finding was that although the older adults' on-line reading times were slower, they did not differ from the young adults in the effect of syntactic ambiguity on word-by-word reading time. In contrast the older adults off-line question comprehension was influenced by the syntactic complexity manipulation in that question comprehension was reliably poorer relative to younger adults for syntactically ambiguous sentences.

Another approach to discourse comprehension involves the study of inference interpretation. Successful discourse comprehension relies on the ability to detect the relationship between two pieces of information and integrate the information with previously stored general world knowledge, in other words draw an inference. Early studies (Cohen, 1979) found that the ability of older adults to draw inferences from spoken messages deteriorated when the messages were presented rapidly, whereas young adults were not affected by the presentation rate. The older adults were also poorer than the young in answering inferential questions after a short passage (Cohen 1979). However subsequent research (Belmore, 1981, Light et al., 1982) failed to observe consistent age differences in inference comprehension. For example, when comprehension was examined on-line, there was no evidence that young and older adults differed in how readily they drew inferences (Burke & Yee, 1984, Light & Alberton, 1988) although they were slower at this task (Hasher & Zacks, 1988). Examination of correlational data suggested that good recall of inferential cues in older adults depended on initial comprehension of the inferences (Till & Walsh, 1980). Age differences were more profound when memory and integration of several parts of a passage with general knowledge was required. A recent study by Grant and Dagenbach (2000) suggested that older adults were less accurate answering questions tapping memory for target inferences and also experienced more difficulty when interpreting unexpected (rather than explicit) inferences. These findings lead to a conclusion that age differences in inference making are most likely to appear when demands on either storage or speed of mental operations are made or when the integration of information with the general knowledge is required.

The last subtest of the TLC-E, Figurative Language, evaluates the ability to interpret metaphors. Metaphor comprehension relies on an ability to extract the similarity between concepts and to abstract relationships. The performance of older adults on tasks that tap this similarity process directly is somewhat variable. When asked to decide whether two words are roughly the same in meaning, to decide whether a word names a category member or whether two words belong to the same category, older adults are no less accurate than the young (Hertzog, Rasind, & Cannon, 1986). These tasks, however, do not require the explicit statement of the basis of similarity between the two objects. Tasks that require more explicit statements of the basis underlying the similarity between objects sometimes yield age differences. For instance, performance on the Wechsler Adult Intelligence Scale (WAIS) Similarities subtest, in which people are asked to explain how two words are alike, declines with age (Salthouse, 1992). Little is known about comprehension of figurative language itself in old age. Some studies have found no indication of age decline. For example, Bayles, Tomoeda & Boone (1988) who tested adults in each decade of life (from their 20s to 70s) on knowledge of common illocutionary (illogical) acts and an ability to relate the meaning of an utterance to the context in which the discourse takes place, found no evidence for age progressive decline. Studies of proverb interpretation, which presumably involves processes similar to those that underlie comprehension of metaphors, indicate on the contrary the presence of age related decline in ability to choose the correct interpretation from a range of choices (Albert, Wolfe, & Lafleche, 1990). The only systematic examination of comprehension of non-literal language was conducted by Light, Owens, Mahoney and La Voie (1993) who found that elderly participants did not differ from the young ones when asked to determine the literal truth or falseness of the metaphoric statements, or when asked to rate the ease of their understanding. Interestingly age differences started to emerge, particularly for the very old group (80+ years) when the participants were asked to make an explicit statement of the meaning of the metaphor.

In summary, the research to date indicates that elderly are likely to experience difficulties on tasks that test language production, in particular the ability to produce syntactically complex constructions. In contrast, the ability of the elderly to perform the tasks that assess comprehension functions (such as resolution of ambiguities, interpretation of inferences and metaphors) declines to a lesser degree. The age-related decline is more likely to emerge on variants of language comprehension tasks that emphasize memorization of information, speed of responding and direct questioning of interpretation.

1.4 Differences Between Young-old and Old-old in Language Functioning

The aged are often treated as a sole cohort and most studies involve comparisons of young with a single group of elderly adults. With current average life span of 75 years and a growing proportion of the population living into their ninth or tenth decade, it is increasingly inappropriate to consider all those over the age of 60-65 as a single cohort. The significance of studying the differences between the young-old (60-74 years) and the old-old (75+ years) is that the extent of cognitive decline becomes more profound with advancing age. Epidemiological studies generally report prevalence rates of moderate and severe dementia at 6-7% of the population over the age of 65 years but the markedly higher rate of 15-20% is found for those over 75 (Kay & Bergman, 1980).

Knowledge of the extent of individual cognitive change varies considerably across different domains of functioning. The research on cognitive functioning over the adult life span indicates that most abilities tend to peak in early middle life, plateau until the late fifties or early sixties and then show decline, initially at a slow pace but accelerating towards the late seventies (Schaie, 1989).

Two types of research methods have generally been used to investigate adult age differences: longitudinal and cross-sectional. A number of extensive longitudinal studies such as the Seattle Longitudinal Study (1956), Berlin Aging Study (1990), and the Kungsholmen Project (1987) have produced information on verbal performance changes in the elderly (see Schaie & Hofer, 2001, for a more detailed review of longitudinal findings). Longitudinal studies have an advantage over cross-sectional research in that they control for cohort effects. The problem with longitudinal research is that it has assessed more general verbal skills (e.g. Vocabulary performance on the WAIS-III) rather than discourse or higher language skills or more detailed experimentally derived measures. Nevertheless, longitudinal studies still provide some valuable insights into performance of different age groups on verbal tasks. For example, the Seattle Longitudinal study examined Verbal Meaning (or ability to comprehend words) over a 28-year period and found that Verbal Meaning continued to increase slightly until the age of 55 years (Bengston & Schaie, 1989). However, by the age of 69 years the decline amounted to 0.33 SD from the initial level. Furthermore, the magnitude of average decline was virtually trivial for the middle-aged cohort (mean age 57 years), modest for the young-old cohort (mean age 71 years) but very substantial for the old-old cohort (mean age 85 years). Interestingly, only 32% of the sample showed substantial decline when they reached old-old age, suggesting that great variability in performance is characteristic of that age group. The only longitudinal research we are aware of that investigated discourse skills in the elderly was conducted by Kemper, Greiner, Marquis, Prenovost and Mitzner (2001). They examined language samples from the Nun Study for grammatical complexity and density of ideas (propositions). Their results indicated that grammatical complexity and idea density gradually decreased across the life span. At a younger age the participants were using sentences containing many embedded and subordinate clauses. They were also able to convey many ideas using few words. By the time the participants reached their late eighties they were using these sort of complex constructions much less frequently. They also tended to use more words to convey the same number of ideas.

Only a very limited number of cross-sectional studies have compared the abilities of young-old and old-old elderly in general, yet alone in terms of language functions. The results obtained from these studies are comparable with the longitudinal data in terms of the old-old elderly showing steeper declines than the young-old. Comprehension of verbal material was found to substantially decline with each decade of life (Van der Linden, Hupert, Feyereisen, et al., 1999). In particular, the 70- 80 year-olds were found to experience significant difficulties in on-line processing of complex syntactic constructions in comparison to the 50-60 year-olds. Specifically, as the structural complexity of the sentences increased so did the differences between the old-old and the young-old (Obler, et al., 1991). The old-old (over the age of 80 years) were also found to be affected to a greater degree by the syntactic complexity of the sentence when memorization of the material was required for successful performance (Kemper & Anagnopoulos, 1993, Zelinski & Miura, 1990). And as previously mentioned comprehension of figurative language was found to be relatively resistant to age-related decline with a decrease in performance not observed until the individual reaches the 80's decade (Light, 1991). The language production abilities of the old-old also suffer in comparison to that of the young-old. The discourse of the old-old was noted as being characterized by an increase in the number of inaccuracies, repetitions, difficulties finding words and tip-of-the-tongue experiences (Heller & Dobbs, 1993).

In summary, the cross-sectional and longitudinal data indicate that old age is characterized by progressive decline in language functioning, which accelerates when the individual reaches 75–80 years of age. Despite of the portrait painted here of the "average old-old adult", it is important to note that there is great variability in individual performance in the old-old age with some individuals maintaining age-constancy well into old age (Hultsch & Dixon, 1983).

1.5 Cognitive Mediators of Language Performance in Older Adults

The overview of language functioning in the elderly suggested a complex pattern of both impairment and sparing of abilities. This pattern of language functioning presents a fundamental challenge to theories of cognitive aging, which must explain why some aspects of the language system are more susceptible to the effects of aging than others. Current theories of language functioning in the elderly have concentrated on implicating processing differences in the intricate activities associated with the working memory system. The literature can be summarized as being guided by three dominant hypotheses 1) the working memory discourse hypothesis (Stine, 1990), b) the inhibitory hypothesis (Hasher & Zacks, 1988) and 3) the cognitive slowing hypothesis (Salthouse, 1980), each of which is summarized below.

1.5.1 *The working memory hypothesis*

The expression "working memory" denotes a complex collection of theoretical constructs that overlap in various ways. The original definition of working memory put forward by Baddeley (1986) refers to system comprised of multiple specialized components of cognition that allow humans to comprehend their immediate environment, retain information about past experiences, and formulate, relate and act on current goals. The specialized components include the central executive (the supervisory system), which is the key factor involved in the control and regulation of working memory. Two specialized temporary memory systems, a phonologically based store (the phonological loop) and a visuospatial store (the visuospatial sketchpad) are subdivisions to the central executive system (further subdivisions of the subsidiary systems might also exist (Baddeley & Logie, 1999)). Working memory is thought to be involved in moment to moment monitoring, processing, maintenance of information in everyday cognition.

Especially important in the context of the current study is the widely accepted view that working memory processes are orchestrated in the service of higher cognition, particularly higher language functions (Miyake & Shah, 1999). A natural extension of this hypothesis is that any age-related decline in language function occurs when there is a requirement for controlled attention and the simultaneous storage and manipulation of complex information, which may be due to a reduction in the efficiency of working memory storage or processing functions (Kemper, 1986, Kemtes and Kemper, 1997, Stine, Cheung & Henderson, 1995). Broadly defined, storage and processing has been labeled as "capacity" and strongly implies a "space" metaphor. The assumption is that the more limited working memory capacity of older adults increases the likelihood that recently processed propositions will be poorly formulated or maintained, and hence inadequately incorporated into text presentation. Consider for example the task of reading. The reader has to recognize individual words, parse the words into phrases and clauses, establish logical and temporal connections among the clauses, determine the referents of the pronouns, and infer unstated clauses and consequences of events. Moreover, the individual must perform all of these operations simultaneously. Limitations in capacity are thought to result in the breakdown in one or more working memory operation leading to slowing or difficulty in language processing.

In agreement with this proposal, evidence has accumulated in recent years that some aspects of working memory might be particularly sensitive to the effects of age (Salthouse, 1991; Van der Linden, Beerten & Pesenti, 1998; Van der Linden, Bredart and Beerten, 1994). The elderly seem to be relatively unimpaired on tasks that call for passive storage of small amounts of material. Rather, age differences emerge when the participants simultaneously store and manipulate information. The classic example is that elderly subjects differed only slightly from younger adults when asked to remember strings of digits or words (regular span tasks). Little or no change on digits forward occurs as the person ages (Nettlebeck & Rabbitt, 1992) and longitudinal data suggest that performance on digits backwards (a task which some authors (e.g. Light, 1990) believe is a sensitive measure of working memory function) remains relatively stable when elderly individuals are retested over 5 year intervals (Colsher & Wallace, 1991). On the contrary, obvious age-related decline is observed when the elderly perform the Daneman and Carpenter (1980) working memory (listening / reading) span task (Wingfield, Stine, Lahar & Aberdeen, 1988; Salthouse, 1991, 1994) (see Table 2). This task requires the participant to listen to, or read, a list of unrelated sentences while answering comprehension questions about the sentences and then recall the terminal word of each sentence (typically, two to five terminal words; the requirement for serial recall may not be important). A meta-analysis that reviewed age differences in working memory indicated that the average effect size for this working memory span task was very large (d=-0.81), placing it at the 21st percentile of the adult age range (from 16.9 to 81.2 years) performance distribution. Furthermore the Daneman and Carpenter span yielded larger age differences than the Digit Span task (d between –0.35 and –0.53 for Digit Span) (Verhaeghen, Marcoen & Goossens, 1993). These findings suggest that the processing resources of the elderly become overtaxed when they are asked to hold information in memory while simultaneously performing simple comprehension operations. In support of this notion, research by Wingfield et al.(1988) as well as other similar findings (Van der Linden, et al., 1994, Van der Linden, et al., 1998) indicate that age-related changes in working memory are characterized by a decline in flexibility and processing abilities of one or more aspect of working memory, whereas more automatic processes, in particular operation of the phonological loop, remain intact.

Table 2. Mean working memory span scores for young and older adults in seven studies on age effects.

	Task	Young	Old
Kemper & Sumner (2001)	L	3.8	2.7*
Light & Anderson (1985)	L	3.6	3.08*
Marmurek (1990)	L	2.78	2.13*
Mc Ginnis & Zelinski (2000)	L	3.29	2.85*
Pratt & Robins (1991)	L	2.7	2.32*
Tun et al. (1991)	R	3.89	2.57*
Stine & Wingfiled (1990)	R	3.33	2.38*

Note: *- significant difference at p<0.05 level, L= listening span, R= reading span

Support for the hypothesis that working memory capacity limitations account for age-related language-processing problems is largely correlational. Working memory span measures have been found to correlate with language-processing measurers such as reading speed for syntactically complex sentences (King & Just, 1991, Kemper, 1986, Obler et al., 1991), the ability to interpret linguistic ambiguities (Miyake, Carpenter & Just, 1994), the ability to assign referents to pronouns (Just & Carpenter, 1980, Light & Capps, 1986), speed of reading (Miyake, Carpenter, & Just, 1994), the ability to perform comprehension operations under extrinsic memory load (Tun, Wingfield & Stine, 1991) and comprehension of sentences presented at a fast rate (Stine & Wingfield, 1987). Based on this evidence, Just and Carpenter (1992) suggested that working memory capacity is particularly necessary for language processing when the processing demands are high. Thus individuals who have limited working memory capacity (as reflected in their low reading or listening span scores) would experience difficulty in processing complex discourse structures. Hence, the elderly as a group would be expected to have mild to moderate complex language and discourse function deficits, derived at least in part from their working memory deficiencies.

This original working memory hypothesis was carefully re-examined by Caplan and Waters (1999), who have considered a number of lines of evidence from studies of young and older adults, as well as individuals with aphasia and dementia. They distinguished between immediate, interpretative, syntactic processing and post-interpretive semantic and pragmatic processing. Caplan and Waters (1999) argue that there is little evidence to support the hypothesis that working memory limitations affect immediate syntactic processes. Instead, they concluded that working memory limitations affect post-interpretative processes involved in retaining information in memory in order to recall it or use it (e.g. answering comprehension questions). In a variety of studies comparing adults divided in to groups (upper and lower quartiles) based on measures of working memory, Caplan and Waters (1996) note that effects of syntactic complexity did not differentially affect high versus low span readers or listeners. They also report that secondary tasks that impose additional processing demands on working memory did not differentially affect the processing of complex sentences on-line. The support for Caplan and Waters model comes from on-line sentence processing studies conducted by Kemtes and Kemper (1996, 1999) and Waters & Caplan (2001) which demonstrated that only post-comprehensive processes (such as question answering) were affected by age and working memory limitations, while the immediate syntactic analysis was not. It should be noted that Caplan and Waters (1999) theory concentrates mainly on comprehension of complex syntax and does not adequately address other aspects of discourse functioning that may decline with age (e.g. changes in production).

There are also other characterizations of working memory apart from limited-capacity storage component of human information processing system; these assume general capacity limitations on the temporary maintenance of information. The contested issue is whether the control of attention is a fundamental attribute of working memory (e.g. Baddeley & Logie, 1999; Engle, et al., 1999) or whether working memory is better characterized as the activation of long-term (LT) memory and LT-working memory (Cowan, 1999). It is highly likely, however, that working memory is a multifaceted phenomenon. Indeed, Baddeley (2000) now proposes a limited capacity "episodic buffer" as an additional element to his traditional tripartite model. This hypothetical episodic buffer has dedicated storage processes, whereas the central executive is left the tasks associated with the control of attention and of the subsidiary working memory systems. The important characteristics of the buffer are that it holds a limited amount of information from a range of modalities and integrates this with information from LT memory, particularly semantic memory. This idea accommodates the activation/storage views of working memory, while still retaining the domain-specific phonological and non-verbal subsidiary systems. Baddeley (1996; 2000) indicates that the Daneman and Carpenter listening span test may tap the role undertaken by the episodic buffer.

Engle et al. (1999), however, presents a different view of the listening span test arguing that the test is the indirect reflection of the capacity for, or the limits of, controlled sustained attention. They argue that individual differences on measures of working memory primarily reflect differences in "working attention" capacity, particularly in situations involving interference and distraction.

To study the controlled-attention capacity of working memory Engle et al. (1999) asked the participants to retrieve in 10 minutes as many exemplars of a given category (e.g. animals) as possible, without repeating previously produced exemplars. The high span subjects (based on the working memory capacity as determined by the arithmetic operations span task) were only slightly better at the task than the low span subject during the first minute, but the level of disparity increased thereafter. In further experiments the participants were asked to generate animal names under two conditions: one with concurrent load of reading out loud digits and one without the load. The attention demanding concurrent task hurt the performance of the high span subjects but had no significant effect on the performance of the low span subjects. Thus, the high span individuals used controlled effortful search to perform retrieval from natural categories, whereas for low span individuals retrieval was more dependent on the automatic, effortless process of spreading activation.

Currently we are aware of only one study (Harris, 1994) that investigated the relationship between working memory functioning and performance on the TLC-E. Harris evaluated the performance of Alzheimer's dementia patients on two working measures (digit lag and digit ordering) and combined the scores on these tasks into a single composite. Strong simple associations between the working memory composite and the TLC-E composite ($r\ (10)= 0.51$, $p<0.01$) were reported by Harris. Of the four subtests comprising the TLC-E only two, Making Inferences and Recreating Sentences were found to have statistically significant correlations to working memory composite (0.53 and 0.47 respectively). No evidence of a significant relationship between working memory and either Ambiguous Sentences ($r\ (10)= 0.35$, n.s.) or Figurative Language ($r\ (10)=0.41$, n.s.) was found. These findings led the author to conclude that the ability to interpret ambiguity and understanding metaphors did not tax the working memory processing resources of the participants. However, it should be noted that the TLC-E scores in Harris's study were presented in relation to the scaled scores for the young children and no control data was provided, which might have negatively affected the reliability of the findings.

So far, precisely how working memory influences complex language skills is unclear. It is assumed that resource limitation is a principal factor, which in turn may be related to issues of controlled attention or problems in the temporary activation of long-term semantic knowledge and recent episodic information. In any case, the provision of several working memory measures is clearly warranted. Another idea suggests that working memory problems in the elderly can be described in terms of inhibitory processes.

1.5.2 *The inhibitory efficiency hypothesis*

This account of older adults language problems has been put forward by Hasher and Zacks (1988, 1999). They proposed that that inhibitory mechanisms weaken with age and permit the intrusion of irrelevant thoughts, personal pre-occupations and idiosyncratic associations during language processing (Hasher & Zacks, 1988). These irrelevant thoughts compete for processing resources, such as working memory capacity, and impair older adults' comprehension and recall. Two studies by Connelly, Hasher and Zacks (1991) that provide key support for the inhibition theory found that older adults reading times and comprehension was substantially impaired by the presence of distracting words or phrases interspersed throughout the text in comparison to the young adults. Age differences consistent with decreased efficiency in inhibitory processing have been found in other studies too, such as stop signal studies (Kramer, Humprey, Larish, Logan & Strayer, 1994), Stroop studies (Houx, Jolles & Vreeling, 1993) and memory for inferences investigations (Hasher & Zacks, 1988). Hasher, Zacks & May (1997) postulated three functions of inhibition: to prevent irrelevant information from entering working memory, to delete irrelevant information from working memory, and to restrain probable responses until their appropriateness can be assessed. They argued that as older adults suffer from deficits in inhibition, their language processing will remain intact on the tasks that do not require inhibitory involvement. But, when inhibitory mechanisms are required to block out distraction, clear away irrelevancies or switch between tasks, deficits would emerge. This hypothesis received support from the study by Kwong See and Ryan (1995) who examined whether age differences in text processing are attributable to working memory, speed or inhibition. Their analysis suggested that older adults' text processing differences could be attributed to slower processing and less efficient inhibition rather than working memory.

However, more recent experimental work appears to cast doubt on the inhibition hypothesis. Gamboz, Russo & Fox (2000) and Grant and Dagenbach (2000) found intact negative priming (a popular measure of inhibition) in their elderly samples, which undermines the fundamental idea of diminished inhibition. More importantly, Grant and Dagenbach (2000) found that negative priming was unrelated to changes observed on their working memory and discourse processing tasks. Similarly, Salthouse and Meinz (1995) found that inhibition as measured by the Stroop task was modestly related to age differences in working memory, yet the relationship between age and working memory remained when the effects of inhibition were removed, unlike the influence of their speed of processing measure.

1.5.3 *The processing speed hypothesis*

The other main rival of the working memory hypothesis is the processing speed hypothesis, which suggests that age-related slowing in processing speed results in deterioration in cognitive functioning in the elderly. Salthouse (1996) suggested two mechanisms through which general slowing may cause errors and disrupt performance. First, some cognitive operations may be executed too slowly for successful completion in the time available, causing an increase in errors. Second, information from different sources may become available to the central processor so slowly that the earlier information has decayed or is no longer active by the time the later information arrives. As a result, cognitive operations that depend on the simultaneous availability of both sources of information can no longer be executed. This slow availability of information to the processor would cause errors even in tasks without time constraint. Hence, as the operations involved in the construction of a discourse representation are time-consuming, age-related slowing is thought to account for age differences in language performance.

The slowing hypothesis enjoys considerable support in the literature. It has been consistently demonstrated that older adults evidence considerable, age-progressive slowing on processing speed tasks (Salthouse, 1991, 1993, Salthouse & Babcock, 1991). It has also been demonstrated that this slowing is related to decline in language functioning. For example, memory for text and inference generation in older adults is specifically impaired by fast presentation (Tun, Wingfield, Stine & Mecsas, 1992). Analysis of reading times also indicates that older adults need more time to integrate ideas that are conveyed in sentences with a higher propositional density (Stine & Hindman, 1994). Older adults were also found to pause more frequently than younger ones to organize new information during reading (Stine, Cheung, and Henderson, 1995).

Salthouse (1991) has argued that an age-related decline in processing speed causes the decline in working memory. He obtained selected measures of general cognitive functioning, as well as measures of processing speed and working memory capacity in younger and older subjects and found that age related differences on general cognitive measures were related to differences in working memory measures. However working memory differences were in turn accounted for largely by differences in the processing speed measures. Salthouse (1991) used this evidence to argue that individual differences in working memory simply reflect a more basic difference in processing speed. Indeed a range of experiments conducted by Salthouse demonstrated that speed of information processing was a primal factor in most cognitive processes such as spatial rotation, matrix reasoning, associative memory tasks, paired associates learning and free recall (Salthouse, 1994, Salthouse, 1993). An obvious extrapolation from these results is that language performance differences may be fundamentally mediated by differences in processing speed. Indeed, Kwong See and Ryan (1995) using a hierarchical regression approach found that when the age differences in speed and inhibition were controlled for, the working memory did not predict language performance (reading comprehension, sentence recognition and text recall). The authors concluded that frequently reported associations between working memory and language are nothing more than a by-product of working memory sensitivity to age differences in speed and inhibition. However, such a strong conclusion may be premature. Kwong See and Ryan study had a number of methodological shortcomings (see Van der Linden, et al., 1999 for a more detailed review), which included poor measurement of both working memory and language constructs. Kwong See and Ryan findings were re-examined by Van der Linden and colleagues (1999). These researchers used a comprehensive range of working memory, speed, inhibition and language comprehension measures, as well as a more sophisticated statistical technique (structural equation modeling) which permitted to tests theoretical causal relationships between the variables. The results from Van der Linden et al. (1999) study clearly indicated that the contribution of speed and inhibition to language comprehension was indirect and was mediated through working memory. Thus, the working memory remained to be a central explanatory principle of older adults' poor performance in language tasks even after control for the contribution of speed and interference.

In summary, currently there are three major ways in which age-related language decline is conceptualized: decline in working memory capacity or controlled attention, decreased ability to resist interference, and slowing of processing speed. It would be a mistake to assume that these theories are totally independent. In fact it is possible that age variance in language performance can accounted for by a combination of these mediating factors.

1.5.4 *Other contributing factors: verbal knowledge and long-term memory*

Apart from working memory, processing speed and inhibition, it is possible that other variables can mediate the performance of elderly on language tasks. The most obvious candidates would be verbal knowledge and long-term memory. With regard to verbal knowledge the idea is that experts compared to novices have a richer, better-organized vocabulary that can be activated in the course of processing which may give the appearance of increased working memory capacity. However, assessments of the contribution of vocabulary suggest that variation in vocabulary scores does not entirely account for differences in working memory span tasks (Daneman & Green, 1986). Vocabulary-based explanations are less useful in accounting for on-line syntactic processing. Partial role of vocabulary is less salient in sentence production and processing tasks in which the structure and vocabulary are familiar (Carpenter, Miyake & Just). The vocabulary hypothesis is also less compatible with the general effects found in aging, given that older adults often have greater vocabulary knowledge, yet show systemic decrements in language tasks that have high processing components (Salthouse, 1980).

Whilst the contribution of vocabulary to age differences in language is thought to be relatively unimportant, long-term memory ability differences may mediate performance to a greater extent. Thus it is possible that person's ability to learn and retain and retrieve information may affect their language functioning. The strength of this possibility is that it has been well established that long-term memory declines with age (Grady & Craik, 2000). This pattern of decline is well reflected in the scores the elderly population obtains on the Delayed Memory Index of the Wechsler's Memory Scale (WMS-III Manual, 1997). It has also been demonstrated that long-term memory ability plays a crucial role in the ability of the elderly to perform language tasks that demand post-comprehensive revisions (Kemper & Kemtes, 2000). When elderly are required to perform comprehension tasks that place heavier emphasis on the delayed memory for what is comprehended the age differences are more likely to emerge than when only immediate comprehension is assessed (Kemper & Kemtes, 2000, Van der Linden, et al.,1999).

Recently, Ericsson and Delaney (1999) have argued that individual differences on the span task of Daneman and Carpenter can primarily be explained in terms of differences in the ability to encode efficiently the presented word in long-term memory along with the appropriate associations to facilitate subsequent recall. In support of their statement they cite a study by Engle, Cantor and Corello (1992) who found that high IQ subjects allocated their resources strategically to the final words presented, as well as evidence from Daneman and Carpenter (1980) that participants with high spans report to have actively engaged in encoding operations. Language functioning and comprehension may also depend on the successful encoding and retrieval of information (Ericsson and Delaney, 1999). For example developmental studies have found that children's ability to comprehend text was related to their encoding skills (Adams, Bell, Perletti, 1995). Long-term storage abilities were also found to mediate the richness of situational descriptions provided by the participants (Post, Green & Bruder, 1982). Decline in LTM was demonstrated to be associated with reduced verbal fluency and increased tip-of-the-tongue experiences in old age (Verhaeghen et al., 1993). Certain discourse tasks (e.g. making inferences) are more likely to be affected by LTM capacity than others. Readers'/listeners' ability to access relevant knowledge during inference comprehension is essential. In particular, the validation of bridging inferences (the type of inference utilized by the TLC-E) depends on the reader/listener retrieval of the knowledge that validates the idea mediating the current sentence and its antecedent. A study by Singer and Ritchot (1996) demonstrated that the individual differences in inference comprehension were related to both differences in working memory and in the ability to retrieve the relevant knowledge. Inference understanding depends on the construction of the situational model to which the text refers. The situational model integrates text ideas and general knowledge in a manner that may not be reflected in the original structure of the text. The construction of an integrated situational model may have a dramatic effect on memory retrieval. Thus, in agreement with Ericsson and Delaney's proposal (1999), the effectiveness of the situational model constructed in working memory will effect the efficiency of retrieval of relevant information from the LTM. In other words the subjects who are high in working memory efficiency would also exhibit superior access to relevant knowledge in LTM (Cantor & Engle, 1993).

In summary, the literature suggests that richness of vocabulary may have little effect on individual differences in language processing especially for the tasks that require manipulation and organization of information. On the other hand the ability to efficiently retrieve information from the LTM may be a marker of fluent comprehension/production process, with working memory capacity making an independent contribution.

1.6 The Present Study

The present study had two broad objectives. The primary aim was to examine the performance of elderly individuals on the Test of Language Competence-Expanded Edition (TLC-E) (Wiig & Secord, 1989) and generate provisional age-appropriate norms if necessary. The second aim was to identify age differences on other variables (e.g. working memory, processing speed, interference and long-term memory) and then to evaluate the contribution of these variables to performance on the TLC-E.

1.6.1 *TLC-E Performance*

First, scores obtained by the two elderly groups (young-old (65-74 years) and old-old (75-89 years)) on the TLC-E were compared to those obtained by a reference group aged 20-34 years. As previously described, the TLC-E incorporates four subtests that assess the ability to resolve ambiguities, make inferences, construct sentences and interpret figurative speech. Due to the uniqueness of the TLC-E as a test of language functioning (see section 1.2 on the TLC-E as a measure of higher language skills) and the lack of data with regard to the performance of adults on this test, it is difficult to make any direct predictions about the magnitude of expected age differences. On one hand, the TLC-E tasks do not heavily rely on memory or speed of responding, the factors that normally compromise the performance of elderly (Kemper & Kemtes, 2000). On the other hand, the TLC-E requires a degree of planing, flexibility and ability to manipulate the information, skills that are, according to Caplan and Waters (1999), impaired in the elderly. Furthermore, some subtests emphasize language production over comprehension and involve greater manipulation of information in order to arrive at an answer (e.g. Recreating Sentences subtest) than others. The *Recreating Sentences* subtest directly tests production of syntactically complex sentences with two or more propositions. Given that current research suggests that older adults tend to exhibit clear deficits in language production (Burke & Mackay, 1997), it was expected that elderly may demonstrate substantial decline on the Recreating Sentences subtest. The magnitude of decline on this subtest was expected to be more profound for the old-old group since syntactic complexity has been shown to decrease progressively across the life span (Kemper & Kemtes, 2000). The *Figurative Language* subtest was not expected to yield large age differences except in the old-old group, as previous research by Light et al. (1993) found a decline in this aspect of language to occur only in the 80+ year-olds. Controversy surrounds the question of whether the elderly experience a deterioration in the ability to resolve ambiguities, with more recent research (Kemper & Kemtes, 2000) suggesting that it is only when the older adults' comprehension is questioned directly that age differences start to emerge. As the *Ambiguous Sent*ences subtest of TLC-E assesses resolution of ambiguity by question and answer rather than on-line technique, mild deficits were expected to be observed in the young-old with potentially increased deficits in the old-old elderly. The readiness with which older adults draw inferences was shown to be affected by such factors as pace of presentation and/or the need to memorize the information (Light, 1990) with little differences found when comprehension was examined immediately after the inference presentation. As the *Making Inferences* subtest of the TLC-E does not require memorization of the target sentences but possibly requires episodic recall of similar social situations, to assist inference making, slight age related deficits were expected to be observed on this subtest.

To reiterate, the expectation was that differential rates of decline would be observed for each of the TLC-E subtests, with age-related deficits more evident on some but not other subtests. It was also expected that the performance of old-old elderly on the TLC-E would mimic the trends currently observed in the verbal functioning of the elderly, in terms of the old-old group sustaining greater loss of function than the young-old.

The present study also aimed at developing provisional norms for the TLC-E for the elderly in the age bands of 65-74, 75-89 years and a young reference group aged 20-34 years. The reference group performance has traditionally been considered in the psychometric literature as reflective of optimal cognitive performance of a healthy adult (Wechsler, 1975). For this reason the reference group's age range was chosen to represent the performance of young adults in the current study. Several considerations guided the selection of age bands for the elderly. The lower limit (65 years) is normally considered as an arbitrary cut-off when differentiating middle adulthood from old age (Schaie, 1989). The upper limit (89 years) was chosen based on the consideration that since increasing number of New Zealanders are living well into the old age (Statistics New Zealand, 1996) the age range should also extend to include the very old. The 65-89 age band was then subdivided into two bands: the young-old and the old-old. This was done in a fashion that approximated the WMS-III age bands. The WMS-III uses 5 year steps to subdivide its normative samples. It was thought that subdivision of the age bands into 5 year steps would not be appropriate for the present purposes, since little cognitive changes are normally observed when individuals are reassessed in 5 year intervals (Schaie, 1989); instead a ten year step was chosen yielding two bands 65-74 years and 75-89 years. This subdivision is also consistent with previously introduced definitions, that "young-old" are thought to represent those in their 60's and early 70's, and the "old-old" are represented by those in the late 70's and beyond (Schaie, 1983).

Currently, the TLC-E has only US norms for ages 5-9 years and 9-18+ years. Normative data from older people would be of value as a useful point of comparison for studies investigating the language ability in other groups, particularly patients suffering from neurodegenerative disorders. For example, it was mentioned earlier that Harris (1994) collected data on patients with Alzheimer's dementia, which indicated that the TLC-E, particularly some of its subtests, would be valuable as a diagnostic tool for this condition. Harris's study included patients with moderate dementia, for which reason she used an easier version (Level 1 for children between 5-9 years) of the test. The Level 2 (adolescent) version was used in the present study, as it was thought to be more suitable for identification of early stages of dementia.

For the purpose of establishing provisional norms an important aim of the study was to recruit a sample of participants who would be representative of the New Zealand population. Given the nature of the cognitive variables under consideration, the use of population-appropriate levels of education in sample selection was regarded as crucial. The majority of studies that investigate language functioning in the elderly tend to recruit the participants with higher levels of education (see Tun & Wingfield, 1993 for a review), mainly due to difficulties associated with participant recruitment. Such samples may not adequately represent the ability level of the general elderly population. The present study obtained a sample in which the proportion of participants with a certain educational qualification approximated that of the national population by gender for each age group. This strategy enhances the representativeness of the sample, and increases the reliability of provisional norms. It also ensures variability in performance, which is essential for detecting associations with other variables (see section 1.6.2). The TLC-E was originally designed to detect problems in language competence, so it is sensitive to errors but has a very truncated upper range, which may pose limitations when used with healthy, well-educated participants. Well educated participants would be more likely to score at the mean or above the mean level producing a restricted range of scores. Recruiting participants of different levels of education assisted in overcoming this problem.

To summarize, the first main aim of the study was to obtain data on the performance of the elderly population on the TLC-E for the purpose of conducting between age comparisons and the development of provisional age-appropriate norms.

1.6.2 *Contribution of other measures to performance on the TLC-E*

The second main aim of the study was to examine the associations between age-related changes in processing skills (working memory capacity, processing speed, inhibitory efficiency and long-term memory) and higher language functions (as assessed by the TLC-E). For this purpose several analytical techniques were used, including analysis of covariance and path analysis. Two main path models were tested and compared: one proposed by Salthouse (1991) where speed is seen as a major mediating factor of age reductions in language, and another proposed by Van der Linden et al. (1999), where working memory plays a crucial role in explaining age differences in language performance and other mediators exert their effects on language through working memory.

Previous research (Kemper, 1986, Kwong See & Ryan, 1995, Light & Capps, 1986) that examined the contribution of processing skills to language function has been criticized for the choice of language measures used (Van der Linden, et al., 1999). Normally, the investigators concentrated on evaluating the contribution of a particular mediator to only one type of language function (e.g. inference making, disambiguation). This strategy, while valuable, has hampered the understanding of how discourse and higher language in general is mediated by limitations in processing resources or other cognitive skills. In the current study the administration of the TLC-E presented a unique opportunity to evaluate the contribution of mediator variables to age differences on discourse skills. The TLC-E combines the scores on four of its subtests to arrive at a single composite. Although varying considerably the subtests that make up the TLC-E all measure the same underlying construct – higher language, discourse skills. Thus, the composite TLC-E score was deemed to provide the most important and more reliable measure of discourse language function, especially in the context of any mediating variables, than any single TLC-E subtest.

Furthermore, the research also aimed at constructing a single factor for each of the mediating variables in question. For this purpose a variety of tasks that purport to measure the relevant constructs were administered. The scores obtained by participants on individual measures were averaged to create composite measures of the relevant constructs. Composite measures have an advantage of increasing the reliability of the statistical techniques used by minimizing the specific variance associated with the single measure and emphasizing the common, construct-related variance.

Because it was assumed, and available evidence seems to be consistent with the assumption that working memory is likely to play a major role in language function (Van der Linden, et al., 1999), the main focus of the present study was on examining the working memory – discourse association. As the characteristics of the working memory system have been subject of a wide variety of theoretical formulations (Shah & Miyake, 1999), the present study selected a number of working memory measures that are potentially related to different aspects of the system, it attempt to better capture the underlying construct. The selected measures also enabled the evaluation of older adults' working memory abilities.

Working memory measures

The first measure was the Daneman and Carpenter (1980) reading span test. The reading span test has been specifically devised to assess working memory from a combination of computational and storage components and is best seen as a reflection of working memory capacity. As previously described, older adults have been shown to perform consistently worse on this span task. More specifically, the performance tends to decrease sharply from young to young-old, but shows little deterioration with further age (Merugo, Fujii, Yamadori, et al., 2000, Waters & Caplan, 2001). The span task also correlates highly with the measures of language comprehension (Daneman & Carpenter, 1980). In the present study two measures of the Daneman and Carpenter test were utilized, the traditional span measure and the total number of words recalled. The total number of words recalled offers an advantage in terms of having considerably more variability than the traditional score (see Kirasic et al., 1996, p.660)

As the Daneman and Carpenter span task contains a dual task element it might be heavily affected by the capacity for divided attention. In addition it is more obviously related to language than any other more general aspect of working memory (Caplan & Waters, 1999) with some researches even suggesting that the span task measures language processing expertise (Cantor & Engle, 1993). Caplan and Waters (1999) recommended that if the goal is to measure the domain free capacity of working memory then a battery of different working memory tasks that differ in the domain specific processes should be administered. Partly for that reason, other working memory measures that comprise the Wechsler Memory Scale – III (WMS-III) working memory subtests as well as the digit span forwards and backwards and semantic fluency tasks were administered.

Engle and colleagues (1999) regarded working memory capacity as controlled sustained attention. This characteristic can presumably be measured by Engle's extension of the semantic fluency task, which requires sustained effortful retrieval of exemplars from semantic memory. Fluency tasks are often prescribed for single minute episodes, with category or letter varying every minute. Engle (1996; see Rosen & Engle, 1994) suggested that when animal fluency for a single category is measured for longer periods, such as 5 to 10 minutes, it is a sensitive index of working memory. High (upper quartile) and low capacity (lower quartile) subjects differed markedly only as time on task progressed. Thus animal fluency over a 5-minute period provided a second working memory measure.

One interpretation of age differences in semantic fluency is that it involves executive control processes in operations like stopping the last search process and initiating the next one (i.e. switching between retrieval categories). In fact, Troyer, Moscovitch & Winocur (1997) have demonstrated that switching was the main determinant of effective performance on the fluency task. Troyer et al. (1997) also found that old adults exhibited a smaller number of switches between semantic clusters during a fixed production interval but the same number of within cluster transitions, suggesting that switching was likely to reflect the frontal lobe based executive processes. Troyer et al (1997) did not consider switching from the perspective of working memory. As mentioned earlier Engle (1996) regards his longer fluency test as a reflection of controlled, effortful retrieval from semantic memory, because individuals with high working memory span differed from low span subjects in the number of clusters recalled. Furthermore, "controlled switching," which is apparently revealed by this fluency test, is itself seen by some theorists (Baddeley, 1996; Stuss, Shallice, Alexander & Picton, 1995) as another unique characteristic of working memory and controlled attention. Hence the present study aimed at evaluating both components of the fluency task, word production over 5 minutes and switching/clustering.

Increased awareness of the importance of working memory as a cognitive construct has seen the inclusion of new measures in the recent revision of the WMS (WMS-III). Given the wide acceptance of the WMS-III as a clinical tool, its working memory subtests were also included here. Another benefit is that the WMS-III working memory index consists of two subtests, which complemented the other working memory measures employed here. For example, the letter-number sequencing subtest assesses working memory by requiring an individual to hold and process information in memory while rapidly switching between encoding strategies. Meta-analytical studies on working memory measures also suggested that Letter/Number Sequencing is sensitive to age progressive decline and yields age effect size comparable to those produced by the Daneman and Carpenter span (Verhaeghen et al., 1993). The Spatial Span subtest of the WMS-III is a more traditional measure, assessing the spatial processing component of working memory. Baddeley (1996) suggested that performance on the Spatial Span type tasks is not only reflective of the workings of the visuospatial sketchpad but also that of the central executive.

The auditory version of the digit span was also included as a measure of working memory. Digit span, especially its backwards version, has traditionally been used as a measure of working memory in research. However, the age effects obtained with the digit span have not always been consistent, some findings indicating that it may be less sensitive to age related decline (Verhaeghen, et al., 1993). Hence, in the present study the Digit Span served as a "control" measure of working memory.

Other measures: processing speed, inhibitory efficiency and long-term memory

In addition to working memory, the contribution of other factors such as processing speed, long-term verbal memory and inhibition to language performance was assessed. The Stroop task was chosen as a measure of inhibitory efficiency. Current theorizing is that the Stroop interference effect arises as a result of automatic activation of irrelevant information (incongruent color word) that needs to be actively inhibited to produce an appropriate task response (naming the ink color of words) (MacLeod, 1991). The effect is highly robust and age sensitive in that older adults have been found to show more Stroop interference (Cohen, Dustman & Bradford, 1984). Additionally the amount of interference was found to remain relatively constant across middle adulthood but then begin to increase in the 60's (Hartley, 1993). Older adults were also found to have greater difficulty ignoring the printed words and attending to the color of the stimulus (West, 1999). These data support the interpretation of the Stroop effect as an indication of less efficient inhibition of irrelevant information in the elderly. Other researchers however have directly related Stroop interference to the functioning of the working memory system (Shallice, 1988). Shallice argued that Stroop performance comes under direct control of the Supervisory Attentional System (SAS) (an analogue of the central executive in Baddeley's (1986) model). The SAS functions to inhibit more automated responses (word reading) in favor of the less automated but goal relevant response (color naming). Hence Stroop interference can also been seen in a broad sense as a measure of working memory functioning.

Processing speed is another factor thought to mediate language performance. Three measures of processing speed were used in the present study: the Digit Symbol Coding subtest of the WAIS-III, color naming and word reading. The color naming and word reading tasks were the ones utilized in control condition of the Stroop task. These tasks can be considered as simple measures of speed and they have previously been used with success to detect age differences (Kwong See & Ryan, 1995, Van der Linden, 1999). The Digit Symbol Coding was recommended by Salthouse (1992) as a good measure of task-independent information processing speed. The Digit Symbol task has previously been demonstrated to be sensitive to age related decline and examination of Digit Symbol performance in young-old and old-old has also yielded significant differences in substitutions (Luszcz, 1992, Schaie, 1989, Sliwinski & Buschke, 1999).

As indicated earlier, it is also highly likely that one's capacity to learn and retain verbal information may be a factor which affects language functioning. The immediate and delayed verbal memory tasks of the WMS-III were administered to test this possibility. The tasks included were Logical Memory I and II and Verbal Pairs Associates I and II. The Logical Memory tests prose recall. Studies that assessed the performance of elderly on this subtest generally found evidence of some of age related decline, which accelerates as the individual reaches the 80-90's decade (Sliwinski & Buschke, 1999). The Verbal Pairs Associates test one's ability to remember semantically unrelated words pairs. The Verbal Pairs subtest may be more sensitive to progressive deterioration with age than the Logical Memory (McCarty, Siegler & Logue, 1982). It is unclear to what degree the long-term memory retention as measured by the WMS-III subtests mediates discourse functioning. So far, the research suggests that the LTM mainly affects comprehension when memorization of large amounts of information is required (Van der Linden et al., 1999). There are, however, some indications (Singer & Ritchot, 1999) that making inferences may rely on the ability to retrieve relevant knowledge from long-term memory.

1.6.3 *Summary*

Two general questions were addressed in the present study: a) Are there reliable changes in average performance of older adults on the TLC-E discourse tasks in comparison to the young? Are these changes large enough to warrant the development of provisional norms for the elderly on the TLC-E? b) What age differences exist in performance on various measures of working memory, speed of processing, inhibition or long-term memory? And to what extent do the age differences on these variables, with a particular focus on working memory, account for the changes in performance on the TLC-E?

2. METHOD

2.1 Participants

Three groups of participants were involved in the study. The participants tested were 20 young adults (20-34 years), 22 young-old adults (65-74 years) and 20 old-old adults (75-89 years). Table 3 summarizes their characteristics by age group. The sample sizes were derived on the basis of statistical power analysis. The final sample (bar two cases missing in two groups) has a 78% chance of detecting an expected large effect size across means ($f=0.40$, $p<0.05$). Large effect sizes are justified on the basis of previous research (Daneman & Merikle, 1996, Verhaenghen et al., 1993). The total sample size has a 99% chance of detecting a large simple correlation ($r=0.50$) and a 68% chance of detecting a medium ($r = 0.30$) correlation. All participants were recruited from the community through personal contact, a Psychology Department subject pool, Pegasus Lions Club and advertisement notices posted on the Psychology Notice Board (see Appendix for the advertisement notice). All participants were reimbursed $30 for travel for the two test sessions.

Participants were excluded from the study if:

a) English was not their first language (spoken at home)
b) They had a history of neurological disease of trauma
c) They had a history of severe medical illness (e.g. heart attack, Type II diabetes, severe migraine).
d) They had a history of psychiatric illness that required hospitalization or a history of depression in the last 6 months.
e) They had a history of alcohol abuse
f) They had a learning disability.
g) They were taking part in a therapeutic trial or were currently taking any medication that might have negatively affected their performance.
h) They had uncorrected vision or hearing impairment or their hearing or vision was severely compromised.

i) Their Mini-Mental Status Score was below the cut-off score of 24 for normal cognitive functioning (Folstein, Folstein & McHugh, 1975).

j) Their Beck Depression Inventory –II Score was above 17, a recommend cut-off score for depression in research studies (Beck, Steer, Brown, 1996)

The aim of the study was to select age-representative groups. To achieve this aim the study selected the participants in a fashion that ensured that the percent of people in a sample with a certain educational qualification approximated that cited by Statistics New Zealand (1996) for New Zealand population by age and gender. Table 4 provides the percentages of participants with certain qualifications in the present sample, and comparison with the national population.

The following educational classification system utilized by Statistics New Zealand (1996) was used as a guide in recruiting participants for the present study.

1) School qualification (For example: school certificate passes, sixth form qualification, higher school qualification, University Bursary Entrance Examination).
2) Vocational qualification (For example: trade certificate, technicians certificate, apprenticeships, national certificate, national diploma, advanced trade certificate bridge certificate, pre vocational certificate).
3) Higher qualification (For example: undergraduate diploma or certificate, New Zealand diploma or certificate, BA, B.Sc., MA, Ph.D., post-graduate diploma).
4) None of the above

As expected from Table 3 the young adults had more formal years of education than the older adults. A one-way ANOVA detected a significant group effect ($F(2, 59)=4.29$, $p<0.05$), with post-hoc Newman Keuls ($p<0.05$) indicating the presence of significant difference only between the young and the old-old. These differences were expected and were a function of sample selection on educational qualifications.

Table 3. Participant characteristics by age group.

		Age (years)		Gender		Years of education*	
Age group	n	M	SD	Male	Female	M	SD
20-34	20	25.45	(2.79)	9	11	7.88	(2.59)
65-74	22	70.36	(3.02)	12	10	6.41	(3.32)
75-89	20	81.25	(4.19)	10	10	5.13	(2.91)

Note: *- years of education were calculated excluding primary school, that is starting from about the age of 10-11 years.

Table 4. Proportions of study sample by educational qualification and gender in comparison to the national distribution (Statistics New Zealand , 1996).

Subjects' Age/Gender	School Qualifications		Vocational Qualifications		Higher Qualifications		None	
	Sample/Stats NZ		Sample/Stats NZ		Sample/Stats NZ		Sample/Stats NZ	
Males 20-34 years	33%	36%	11%	18%	22%	18%	33%	27%
Females 20-34 years	36%	36%	9%	9%	27%	27%	27%	27%
Males 65-74 years	17%	18%	17%	18%	16%	9%	50%	55%
Females 65-74 years	20%	18%	0%	approx. 0%	10%	9%	70%	72%
Males 75-89 years	20%	18%	10%	18%	9%	9%	60%	55%
Females 75-89 years	10%	18%	0%	approx. 0%	9%	9%	80%	72%

The participants were administered a number of screening measures (see Materials for detailed description) including: Mini-Mental Status Examination (MMSE) (Folstein, Folstein & McHugh, 1975), Beck's Depression Inventory-II (BDI) (Beck, Steer, Brown, 1996), The National Adult Reading Test- Second Edition (NART) (Nelson, 1991), and the 2-subtest IQ from the Wechsler Abbreviated Intelligence Scale (WASI) (Wechsler, 1999) (see Table 5 for group means). Unless otherwise specified all group differences were examined by using one-way ANOVA followed by post-hoc Newman-Keuls ($p<0.05$) comparison. The age groups did not differ in their level of cognitive functioning as measured by the MMSE ($F(2, 59)=1.95$ n.s.) or on their scores obtained on the Beck's Depression Inventory ($F(2, 59)=0.39$, n.s.), a measure for detecting depression in normal population. The intellectual functioning of the groups was evaluated on the 2 subtests (Vocabulary and Matrix Reasoning) of the WASI which generated no significant age effects ($F(2, 59)=1.31$, n.s). Additionally the groups also did not differ in their verbal knowledge ($F(2, 59)= 0.98$, n.s.) as evaluated by the Vocabulary subtest of the WASI. There were however clear group differences detected on the Matrix Reasoning subtest with the young adults performing better than both elderly groups who also differed from each other ($F(2,59)=30.00$, $p<0.01$). The Matrix Reasoning subtest has different discontinue points for older individuals which ultimately can result in different scores obtained. However examination of individual scores suggested that only one older adult could have potentially obtained a higher score if the subtest was not discontinued. In other words, the raw score group differences obtained on Martix Reasoning were likely to reflect real age differences on this task of abstract reasoning. The presence of age-related decline on Matrix Reasoning but not on the Vocabulary subtest is a frequently observed phenomenon in aging research and is indicative of increased vulnerability of fluid intelligence functioning to aging in comparison to the crystallized abilities (Shaie, 1989). On the National Adult Reading Test –Second Edition, a test used for estimation of premorbid intellectual functioning, no significant group differences were obtained on the error score ($F(2, 59)=1.46$, n.s.). Similarly, no group differences were obtained on the NART predicted IQ score ($F(2, 59)=1.4$, n.s.).

Table 5. Group means and (standard deviations) on the screening tests.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)
BDI	4.60 (4.31)	5.59(3.54)	5.45(3.80)
MMSE	28.95 (1.00)	28.77 (1.41)	28.15 (1.57)
NART	18.95 (7.54)	15.4 (7.30)	18.95 (8.51)
NART IQ	107.50 (9.68)	111.73 (9.17)	107.25 (10.68)
WASI IQ	109.65 (12.65)	115.23 (13.42)	109.55 (13.11)
Vocab	60.55 (9.21)	63.32 (9.14)	59.30 (10.28)
Matrix	28.30 (3.11)	22.54 (6.01)	15.45 (5.87)

Note: BDI= Beck's Depression Inventory, MMSE= Mini-Mental Status Examination, NART= National Adult reading Test (2nd edition), NART IQ= an estimated premorbid IQ derived from the NART error score, WASI IQ= 2 subtest IQ form the Wechsler Abbreviated Adult Intelligence Scale, Vocab= Vocabulary raw score from the WASI, Matrix = Matrix Reasoning raw score from the WASI.

Furthermore a series of dependent means t-tests ascertained that there was an absence of significant differences between predicted (NART) and obtained (WASI) IQ for young ($t(19)=1.12$, n.s), young-old ($t\ (21)=1.56$, n.s.), and old-old adults ($t\ (19)=1.07$, n.s.). Thus it is unlikely that the participants in the present study have experienced substantial decline in their intellectual functioning from the estimated premorbid level.

2.2 Procedure

The tests were administered individually to each participant over the course of two 2- 2.5 hour sessions. The sessions were separated by one week. The participants had an opportunity to have short breaks (including imposed short breaks) during the sessions to prevent fatigue. All the participants took part in all of the tests. Prior to the commencement of testing the participants read an Information Sheet that outlined the goals of the study and signed a consent form (see Appendix for Information Sheet and Consent Form). A short screening interview was then conducted by the experimenter, to ascertain relevant details and general information regarding participants' health (see Appendix for the Questionnaire Sheet). The order of test administration was as follows:

SESSION 1: Mini-Mental Status Examination, Daneman and Carpenter reading span test, (short break), Test of Language Competence- Expanded Edition Level 2, (short break), Semantic Fluency Test, Beck's Depression Inventory- II.

SESSION 2: Logical Memory I, Verbal Paired Associates I, Letter-Number Sequencing, Spatial Span, Digit Span (short break, time restrictions permitting), WASI Matrix Reasoning subtest, Digit Symbol Coding, Logical Memory II, Verbal Paired Associates II, (short break), WASI Vocabulary subtest, a Stroop test, The National Adult Reading Test –2nd Edition.

2.3 Materials

2.3.1 *Screening Measures*

A number of additional measures were incorporated into the test battery to provide background information on participants and screen for any possible general deficits.

Vocabulary and Matrix Reasoning subtests

Vocabulary and Matrix Reasoning are the subtests included in the Wechsler's Abbreviated Intelligence Scale (Wechsler, 1999). Administration of these two subtests provided a quick way of estimating person's current IQ. The Vocabulary subtest is a measure of verbal ability and verbal comprehension requiring the participant to give definitions to various more or less common words. Matrix Reasoning is a measure of visual information processing and abstract reasoning skills. Each item in the subtest consists of a stimulus matrix from which a section is missing and five response choices. The participants have to select the completing response from the choices provided. The instructions and scoring procedures for the subtest were conducted as per the manual.

National Reading test -2nd Edition (NART 2)

The NART 2 (Nelson, 1991) is used to estimate premorbid intellectual ability as well as reading ability. The participant was presented with a list of 50 words of increasing difficulty. All words included were of irregular pronunciation and as the participant read down the list errors of pronunciation were recorded.

Mini-Mental Status Examination (MMSE)

The MMSE (Folstein, Folstein & McHugh, 1975) is a common summary screening measure for dementia. The MMSE is comprised of a variety of items that test participant's orientation to time and place, attention, short-term memory, naming, following commands, writing and copying. A score of 24 out of 30 on the MMSE was used as a cut off for normal cognitive functioning.

Beck Depression Inventory- Second Edition (BDI-II)

The BDI is used for detecting possible depression in normal population (Beck, Steer, Brown, 1996) and is based on the typical descriptive statements regarding the symptoms of depression. As depression can adversely affect one's concentration and memory the BDI was utilized to screen for possible depression in the present sample. The score of 17 on the BDI was used as a cut off, as per recommendations provided by Beck et al. (1996) for research purposes. On the BDI the conservative threshold for detecting depression is a score of 14 where 14-19 indicates mild symptomatology, 20-28- moderate and 29-63- severe. To err on the side of caution, feedback was given to the participant if they produced a score above 13 on the BDI. The researcher first talked with the participant explaining the meaning of the score obtained and emphasizing that it was in the participant's best interests to contact their GP to ask for further advice. An additional statement was given to the participant, for them to read, to fill in as they saw fit and sign (plus the researcher gave the participant a copy to retain) that insured the participant's understanding of importance of getting further advice/ evaluation should they choose (see Appendix for statement example).

2.3.2 *Measures of higher language functioning*

Test of Language Competence-Expanded Edition, Level 2

The Test of Language Competence -Expanded Edition-Level 2 (TLC-E; Wiig & Secord, 1989) was designed to identify young people aged 9-18+ as well as adults with language disabilities. The test incorporates an assessment of semantics, syntax, and especially pragmatics, using formats that probe divergent production, cognitive-linguistic flexibility and planning for production. The test comprises four subtests each of which tap a distinct domain of discourse. For each subtest, instructions were provided to the participants as per the test manual. The scoring of the subtests was conducted according to the manual guidelines.

Ambiguous Sentences

The first subtest, Ambiguous Sentences, evaluates the ability to recognize and interpret alternative meanings of selected lexical and structural ambiguities. For example the participant had to provide two interpretations of a spoken sentence that contained an ambiguity (for e.g. *"I've always known that flying planes can be dangerous*"). The sentences were first read to the participants and then presented in a written format.

Making Inferences

The second subtest, Making Inferences, involves the ability to make an inference given two casually connected events. Each item in the subtest describes an event chain in which one or more casual links are missing. One proposition outlines a situational event or script (e.g. *" Jack went to a Mexican restaurant*") and a second phrase presents an outcome (e.g. *" He left without giving a tip "*). The participant's task was to choose from four expressions the two that best reflected the possible intervening events (for e.g. "*He only had enough money to pay for the meal"* or *" He was dissatisfied with the service"* but not *"The restaurant closed when he arrived"* or *"The food and service were excellent"*). The participant had to interpret the proposition, recognize and generate underlying scripts and make global inferences based on his or her knowledge of a possible causal event in the appropriate script. The propositions were first read to the participant and then presented in a written format together with the choices.

Recreating Sentences

The third subtest, Recreating Sentences, evaluates the ability to plan and formulate expressions of intent (speech acts) incorporating key words related to a situation or context. Each set of stimulus words includes a grammatical marker (linguistic concept) commonly used to join two independent sentences in to a complex sentence structure (compound/complex). The remaining words represent vocabulary associated with the noun case, the verb case, the adverbial case, and the modifier. The participant was presented with three stimulus words (for e.g. *neither, week, were*) and an illustration of a prescribed context (e.g. " In the supermarket"). The participant had to construct a grammatically correct sentence using all the given words appropriately.

Figurative Language

The fourth subtest, Figurative Language, evaluates the subject's ability to interpret figurative expressions. The items are all commonly used metaphors or idioms (e.g. *"She seems to be holding all the aces"*). The task required the participant to differentiate between metaphoric and literal expressions, to state the meaning of the expression in the participant's own words, and then to identify an alternative expression for the phrase from the four choices provided (for e.g. *"The odds certainly favor her"* but not *"She is a real card shark"*, or *"She has four aces in her hand"* or *"The chips seem to be down for her"*).

2.3.3 *Working Memory Measures*

Daneman and Carpenter Reading Span Test

The Daneman and Carpenter (1980) reading span test requires the participant to read out loud a list of sentences. The sentences were presented as a set ranging from 2 to 6 sentences in each set. At the end of each set the participant was required to recall from memory the last word from each sentence. To ensure that the participants comprehended the full sentences and not just concentrated on the final words, they were required to indicate whether or not each sentence was true or false immediately after it has been presented. Half of the sentences were true and half false. Prior to testing, 5 practice items at 2 sentences a set length were administered to ensure that the participants understood the instructions and also to permit the estimation of how long the participants took to read each sentence. Subsequently during the main part of testing it was ensured that the same speed of reading was maintained, in order to prevent the participants from taking extra time to rehearse the last words of the sentences.

The span test consisted of 60 different sentences (plus 10 practice sentences), each about 9-16 words in length and ending with a different word. There were three presentations of each set length and the participants were not warned in advance when the number of sentences in a set was about to increase. The sentences were provided by Meredith Daneman (personal correspondence) (see Appendix for sentence examples and administration instructions). The sentences (font size 32pt., Times New Roman) were displayed on laminated white cards (7 x 23 cm) one at a time and centered. Immediately after the participant read the sentence and answered true/false, the experimenter turned over the card to present the next sentence; the presentation of a blank card signaled the time to recall the last word of each sentence. The participants were instructed to attempt to recall the words in the order they were presented and specifically avoid beginning with the last word of the last sentence unless it was the only word remembered.

Irrespective of their performance the participants were administered all of the 60 test sentences. Two measures of reading span were obtained. The traditional span score was calculated according to the instructions provided by Daneman and Green (1986). The maximum set size at which the participant was correct on two out of three in a set was taken as a measure of his or her reading span. Half credit was given if the participant was correct on one set at a particular level. For example if the participant was correct on two out of three three-sentence sets in the reading span the assigned reading span score was 3 and if correct only on one of the three the assigned span was 2.5. The second measure was simply the total number of final words recalled correctly (out of maximum number of 60). The participant's correct number of true/false responses was also recorded.

Semantic Fluency Task

The semantic fluency task was adapted from that described by Engle (1996) and Rosen and Engle (1994). The participants are asked to recall as many exemplars of the category 'animals' as possible in a 5 minute time periods avoiding any repetitions (see Appendix for instructions). Their answers were audiotaped for later scoring. A computer program was developed in the Psychology Department, University of Canterbury, to increase the accuracy of scoring. The experimenter listened to the recording and pressed the "space bar" every time a word was produced, to record the exact time position of each word. Upon completion of the recording, every word produced was entered individually into the computer. Scoring was conducted according to guidelines adapted from in Troyer, et al. (1997):

1) Total number of exemplars recalled in five minutes (minus errors and repetitions).
2) Total number of exemplars recalled in each minute (minus errors and repetitions).

3) Total number of between category switches in five minutes.
4) Total number of between category switches for each individual minute.
5) Mean cluster size for 5 minutes (including repetitions)
6) Mean cluster size for each individual minute (including repetitions).

Clusters were defined as groups of successively retrieved words belonging to the same semantic category, such as farm animals, African animals, New Zealand native animals, etc., and various zoological categories, such as birds, primates, insects etc. The determination of potential semantic categories as listed in the Appendix, was derived from the actual pattern of words generated by the participants during the test performance, using categories developed by Troyer et al. (1997) as a guideline. Scoring was undertaken after all the participants have been tested and a comprehensive list of animal names was developed. During scoring any words produce by the participants that were not included in the developed list thus far were added to the list. Cluster size was counted beginning with the second word in each cluster and then the mean cluster size was calculated. Switches were calculated as the number of transitions between clusters, including single words as per Troyer et al. (1997). Errors and repetitions were included in calculations of cluster size and switches because these were thought to provide information about the underlying cognitive processes. In cases where two categories overlapped with some items belonging to both categories, some items belonging exclusively to the first category and some items belonging exclusively to the second category, the overlapping items were assigned to both categories. For example: for *dog, cat , tiger, lion,* the first two items were scored as *pets*, the last three items were scored as *feline.* In cases where smaller clusters were embedded within the larger ones, or two categories overlapped but all items could be correctly assigned to that same category, only the larger common category was used. For example: *crocodile, alligator, snapper, dori*, are all water animals, but an additional cluster (*fish*) was not scored for the last two items.

As 5 minutes of semantic fluency task was administered in the present study, minute-by-minute rules and five-minutes rules were also developed.

Minute by minute rules: Each minute was considered separate and the calculations were unaffected by what happened before or what would happen in the next minute. This rule applied to the number of words produced, clusters and switches. With regard to cluster size a new count began when each minute started. If the person was going through one category when one minute expired, than the words produced in that minute formed their own cluster; all the words produced in the next minute formed another cluster. For example: *cat, dog* (first minute ends) *canary, rabbit.* All these exemplars are *pets* but only the first two exemplars were included in the cluster size calculation for the first minute. In the same vein the number of switches was also calculated for each minute individually and the calculations were unaffected by what words were produced in the proceeding or following minutes.

Five-minute rules: According to the five minute rule all the words produced across the five minutes constituted the total word number (minus errors and repetitions). The mean cluster size for five minutes was calculated across all the minutes (i.e. as a single period of time), disregarding the time at which each cluster started or finished, with the same rule applying to switching.

Wechsler Memory Scale-III working memory subtests

The two subtests of the WMS –III (Wechsler, 1997) that are used to assess working memory (the Letter-Number Sequencing and the Spatial Span) and also the optional Digit Span subtest were administered. The scores on the Letter/Number Sequencing and Spatial Span subtests were also combined to produce a Working Memory Index. For each subtest manual guidelines were followed when instructing participants or scoring tests.

<u>Letter-Number Sequencing</u>

The letter-number sequencing subtest is a measure of auditory working memory and requires participants to order sequences of numbers and letters, aurally presented in a specified random order and with increasing length of letter/number strings. Participants must report numbers then letters, rearranging the numbers into ascending order and letters into alphabetical order.

Spatial Span

Spatial span is a visual variation of the digit span and has two components, Spatial Span Forwards and Spatial Span Backwards. The examiner presented the participants with the Spatial Span by tapping the prearranged blocks in the specified random order specified by the manual. Participants then had to repeat the tapping sequences either in the order presented (for the Spatial Span Forwards) or in reverse order (for Spatial Span Backwards).

Digit Span

Digit span consists of two components: Digit Span Forwards and Digit Span Backwards. The test required the participant to repeat strings of digits of increasing length that were aurally presented by the examiner, either in the order presented or in the reverse order as per the manual.

2.3.4 *Long-term Verbal Memory Measures*

Wechsler Memory Scale-III Auditory memory subtests

Four subtests of the WMS-III, Logical Memory I and II and Verbal Paired Associates I and II, provided information regarding participants' the ability to recall auditory presented material immediately and after a delay. The scores on these subtests were combined to produce Auditory Immediate and Delayed Indices. Manual guidelines were followed when instructing participants or scoring tests.

Logical Memory I and II

This subtest consists of two short stories. The stories were presented aurally to the participant and after hearing each story the participant had to recall the story as close to the text as possible. For Logical Memory II the participant had to recall the two stories spontaneously after a 25-35-minute filled delay.

Verbal Paired Associates I and II

In the Verbal Paired Associates-I (VPA-I) a list of unrelated word-pairs was first read to the participant. After a 5 second delay the examiner read the first word in the pair and the examinee's task was to provide the second word in the pair. Four learning trials were administered in the VPA-I. The VPA-II was administered after a 25-35 minute filled delay and involved only one trail where the examiner provided the first word in the pair and the examinee had to recall the second word in the pair.

2.3.5 *Inhibitory Efficiency Measures*

Stroop Color-Word task

The capacity to inhibit irrelevant information is often measured through the variants of the Stroop color-word task. In the present study the Golden (1978) version of the task was used. Manual guidelines were followed when instructing participants and scoring the test.

The Stroop test consisted of three pages. The first page listed a hundred color words printed in black ink (i.e. *red, green, blue*), the second page contained a hundred sets of juxtaposed X's (e.g. XXXX) printed in red, green or blue ink, the third page consisted of a hundred color words printed in incongruent ink colors (for e.g. the word *red* printed in green ink). The participants were presented with one card at a time in the following order: words, colored X's, color-words. The participants were instructed to read the words (or name the colors for the X's and color-word conditions) going from top to bottom of each of 5 columns. They were asked to do that as accurately and as quickly as possible. The participants were instructed that in case they make an error the examiner would say "No", they would then have to stop, correct themselves and keep going. The participants were allowed 45 seconds to complete each page with the second page provided if needed. The number of words read/colors named in that time served as a measure of their reading/color naming speed. The participants' errors were not included in the calculations as participants were already penalized for their mistakes by loosing time in stopping and correcting errors.

Four scores were derived on the Stroop task: number of words read, number of colored X's named, number of color-words named and an interference score. As per the formula provided in the manual an interference score was calculated from the color-word card, corrected for the speed score in the following manner: (color word – (word x color)/(word +color))=interference (note, that raw scores were used in interference calculations).

2.3.6 *Processing Speed Measures*

Word reading and Color naming

Word reading and color naming that comprise the control conditions in the Stroop task were also used as processing speed measures. These measures are already described above.

Digit Symbol

The Digit Symbol Coding subtest from the Wechsler's Adult Intelligence Scale – III (Wechsler, 1997) was a third measure of processing speed. The instructions to the participants were provided as per the manual. The participant was presented with a key comprised of a series of numbers from 1 to 9 each having a corresponding symbol. The participant was also provided with stimulus items, which contained only numbers in random order without the symbols. The task of the participant was to pair each number with its symbol by using the provided key in a time of 120 seconds. The correct number of digit substitutions made in the specified time served as measure of processing speed. The obtained raw scores were also converted into the age-appropriate scaled score from the WAIS-III tables.

3. RESULTS

3.1 Statistical Analysis Employed

The data in the present study were analyzed using the Statistica 6 (2001) package. All between group comparisons were conducted using one-way ANOVA (unless otherwise specified), followed by the post-hoc Newman-Keuls ($p<0.05$) comparison (for simple main effects) tests where appropriate. The difference between obtained means and the population means on standardized tests was examined using two-tailed t-tests. All correlation coefficients were Pearson product-moment correlations. A significance level of $p<0.02$ (criterion $r = 0.30$, $df = 60$) was used for correlation coefficients to partially control for multiplicity of tests. Reliability estimates (Cronbach's α) for composite measures were calculated using the Reliability Analysis Statistica module. Path analysis technique was used to evaluate the contribution of mediator variables to age differences in language and to test the theoretically postulated relationships between variables. Structural equation modeling, which estimates the direct and indirect relations among latent variables was not used, as this technique demands large sample sizes (at least 100-150) to maintain the accuracy of estimates (Schumaker & Lomax, 1996)

3.2 TLC-E Performance

3.2.1. *Age group comparisons using raw scores*

This section compares the performance of three age groups on the TLC-E, both in terms of individual subtests performance and TLC-E composite. The composite score was calculated using the method described by Wiig and Secord (1989), where raw scores on all of the four subtests (Ambiguous Sentences, Making Inferences, Recreating Sentences and Figurative Language) were added. Wiig & Secord (1989) point out that although varying considerably the subtests that make up the TLC-E are all measures of the same underlying construct, language competence, and hence can be combined into a single measure. Table 6 presents correlations between the TLC-E measures obtained in the present study. The correlations between the subtests were moderate to high (0.42 to

0.65), which was indicative of a good interrelationship between the measures. There was a high reliability estimate for this composite (Cronbach's alpha =0.79) (acceptable level $\alpha>0.70$; Barrett, 2002, personal communication). Examination of item-total statistics suggested that exclusion of any of the four subtests from the composite would not result in substantial improvements in α levels.

Descriptive statistics for the raw score measures in the four of the TLC-E subtests and the composite is presented in Table 7. Significant group effects were detected for the TLC-E composite score, Ambiguous Sentences, Making Inferences and Recreating Sentences subtests (for F ratios and p values see Table 7). On each of these measures the young adults obtained significantly higher scores than both elderly groups, who in turn also significantly differed from each other. The overall group difference on the Figurative Language subtest just failed to reach significance. No change in Figurative Language scores was observed for the young-old group in comparison to the young. The old-old group, however, obtained significantly lower Figurative Language scores that the young-old and the difference between the young and old-old adults' scores on this subtest just failed to reach significance ($p<0.07$).

The performance of elderly on the TLC-E showed that higher language functions decline with age. Progressive decline in both older age groups was clearly evident on three of the subtests but Figurative Language showed a relatively milder decline that was apparent only in the old-old group.

3.2.2 *Age group comparisons using standard scores*

Currently the TLC-E manual provides standardized scores (M=10, SD=3 for subtest scores and M=100, SD=15 for the composite score) for children up to the age of 18 years and 11 months. The scores obtained by the participants in the present study were first scaled relative to the norms available for the 18y.11m year old group provided by the manual (see Table 8 for scaled scores).

Table 6. Intercorrelations between the TLC-E measures

	1	2	3	4	5
1. TLC-E composite	--				
2. Ambiguous Sentences	.83	--			
3. Making Inferences	.62	.56	--		
4. Recreating Sentences	.78	.65	.62	--	
5. Figurative Language	.54	.46	.46	.42	--

Note: all *p*s<0.02

Table 7. Means and variability for the TLC-E raw scores for the three age groups and for the 'young adults' provided by the TLC-E manual.

TLC-E Measures	TLC-E scores (17.11-18.11y) *n*= 112	Young (20-34y) *n*=20	Young-old (65-74y) *n*=22	Old-old (75-89y) *n*=20	F (2, 59)	*p*
Composite score						
M	167.3	165.25*°	148.50	134.80†	18.11	.0001
SD	13.8	14.13	11.41	21.31		
Median	--	170	150	140		
Range	--	117-181	125-169	97-164		
SEM	6.6	3.16	2.43	4.77		
Ambiguous Sentences						
M	32.6	35.05*°	30.64	26.90†	14.81	.0001
SD	5.1	2.65	4.07	6.67		
Median	--	35.00	31.00	30.00		
Range	--	30-39	20-37	12-34		
SEM	1.5	0.59	0.87	1.49		
Making Inferences						
M	32.8	31.00*°	27.04	23.70†	16.81	.0001
SD	5.1	2.73	3.54	5.29		
Median	--	32	28.00	23.50		
Range	--	25-36	18-32	13-24		
SEM	2.0	0.61	0.75	1.18		
Recreating Sentences						
M	71.2	71.40*°	60.05	55.85†	31.45	.0001
SD	6.0	4.30	5.56	8.70		
Median	--	71.5	60.00	58.00		
Range	--	59-78	50-70	36-70		
SEM	1.6	0.96	1.89	1.94		
Figurative Language						
M	30.7	30.90	30.86	28.35†	3.13	.051
SD	4.9	3.68	3.44	4.04		
Median	--	32.00	32.00	28.50		
Range	--	21-36	20-36	23-36		
SEM	1.6	0.82	0.73	0.90		

Note: *F* ratios and *p* values are for one-way ANOVA on 3 sample groups.

* - Newman-Keuls significant difference between young and young-old ($p<0.05$)

°-Newman-Keuls significant difference between young and old-old ($p<0.05$)

†- Newman-Keuls significant difference between young-old and old-old ($p<0.05$)

When group differences were analyzed using these scaled scores, which more directly reflect scores relative to a normal distribution, the same results as reported for the raw scores were obtained. The only difference was that the overall group effects on the Figurative Language subtest reached significance, with the old-old versus young group comparison also now becoming significant.

More importantly, these scaled scores also enabled the comparison of the present data with that provided in the TLC-E manual. As expected the reference group in the present study performed at a generally comparable level to the TLC-E 18-year-old population. No significant differences were detected between the TLC-E population mean and reference group mean on the TLC-E composite ($p = .45$), the Recreating Sentences ($p = .38$), and the Figurative Language ($p = .74$) scores. However in the current study the young adults obtained on average a higher score (1.2 scaled points) on the Ambiguous Sentences subtest ($p < .05$) and a lower score (1.7 scaled points) on the Making Inferences subtest ($p < .05$).

3.2.3 *Development of norms for the TLC-E*

The small but detectable differences between the reference group performance in the present study and the TLC-E 18-year-old population performance justified the development of separate norms for the older reference group. Moreover, the clear age differences found between the three age groups in the present study emphasized the need to establish provisional norms for the ages 65-74 and 74-89 years on the TLC-E. However, prior to the development of norms, the contribution of variables such as gender, IQ and years of education to age differences on the TLC-E was assessed in order to investigate the need for subdivision of norms according to these variables.

<u>Gender</u>

Table 9 presents mean performance raw scores on the TLC-E by age groups and gender.

Table 8. Mean group differences and (standard deviations) on the TLC-E standardized scores.

TLC-E Measure	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F (2, 59)	p
Composite	98.50 (8.85)	81.86 (7.92)	74.90 (9.89)	37.28	0.001
Ambiguous Sent.	11.20 (1.70)	8.86 (1.88)	7.35 (2.64)	16.93	0.001
Making Inferences	8.30 (2.23)	5.50 (1.59)	4.50 (2.09)	19.93	0.001
Recreating Sent.	9.55 (2.28)	5.36 (1.50)	4.70 (1.45)	44.23	0.001
Figurative Lang.	10.20 (2.69)	9.64 (2.17)	8.30 (2.27)	3.37	0.04

Note: F ratios and p values are for one-way ANOVA.

Table 9. Group means (standard deviations) for males and females on the TLC-E raw subtest and composite scores.

TLC-E	Young (20-34y) M	F	Young-old (65-74y) M	F	Old-old (75-89y) M	F	F (2,56)	p
Composite								
	170.44 (6.31)	161.00 (17.41)	149.83 (10.52)	146.90 (12.78)	133.50 (25.95)	136.20 (16.55)	0.71	0.50
Ambiguous Sentences								
	35.78 (1.20)	34.45 (3.36)	31.17 (3.13)	30.00 (5.18)	26.80 (8.61)	27.00 (4.45)	0.38	0.54
Making Inferences								
	31.00 (1.41)	31.00 (3.55)	27.25 (3.57)	26.80 (3.68)	25.70 (4.69)	21.70 (5.31)	1.58	0.21
Recreating Sentences								
	71.00 (5.65)	71.72 (3.04)	60.25 (6.50)	59.80 (4.52)	57.50 (8.91)	54.20 (8.64)	0.50	0.60
Figurative Language								
	32.44 (3.13)	29.64 (3.75)	31.58 (2.78)	30.00 (4.08)	28.00 (3.46)	28.70 (4.72)	1.61	0.32

Note: M= males, F= females. F ratios and p values cited are for ANOVA group by gender interaction.

The data presented in Table 9 suggest an absence of substantial gender differences which was confirmed by a series of *t*- tests that detected non-significant differences between the scores of males and females in each age group on each individual subtest. Similarly a 3 (age group) x 2 (gender) ANOVA conducted on each of the TLC-E measure indicated an absence of significant main effect for gender as well as an absence of significant group by gender interaction (see Table 9 for *F* ratios and *p* values). Thus the examination of scores of males and females on the TLC-E indicated an absence of gender differences for these measures.

IQ and education

To explore the contribution of IQ and education to group differences on the TLC-E an approach recommended by Clegg & Warrington (2001) was used. Firstly, the correlations between the TLC-E scores, IQ, years of education and age were computed (see Table 10). TLC-E measures correlated highly and significantly with age (apart from the Figurative Language subtest), weaker correlations between the TLC-E measures and IQ and education years were detected, which were substantially and significantly lower ($p<0.05$) (as evaluated using the test for differences between two correlation coefficients) than those between the TLC-E subtests and age. Two other points should be noted about the observed correlations. Firstly, the Figurative Language subtest demonstrated stronger associations with IQ than age or education. Secondly, the Recreating Sentences performance was more strongly related to years of education than any other of the TLC-E subtests.

To further clarify the contribution of years of education and IQ to observed group differences on the TLC-E these variables were entered as covariates in ANCOVA's of age effects for each measure. The addition of years of education as a covariate did not markedly affect the size of the observed group effects: the TLC-E composite ($F(2, 58)= 13.45$, $p<0.001$), Ambiguous Sentences ($F(2, 58)= 11.00$, $p<0.001$), Making Inferences ($F(2, 58)= 13.07$, $p<0.001$) and Recreating Sentences ($F(2, 58)= 23.88$, $p<0.001$). Similarly, when IQ was entered as a covariate the significant group effects were also maintained for the TLC-E composite ($F(2, 58)= 21.35$, $p<0.001$), Ambiguous Sentences ($F(2, 58)= 16.77$, $p<0.001$), Making Inferences ($F(2, 58)= 20.35$, $p<0.001$) and

Recreating Sentences ($F(2, 58)= 39.92$, $p<0.001$) subtests. The addition of years of education or IQ as a covariate reduced the *F* ratio ($F(2, 58)=1.97$, n.s.) and ($F(2, 58)= 2.19$, n.s.) respectively for the Figurative Language.

In summary, gender, IQ and years of education had minimal contribution to age differences on the TLC-E. Age showed a strong negative correlation with three out of four TLC-E subtests, particularly with the Recreating Sentences subtest. The Figurative Language subtest was the only TLC-E subtest that was poorly correlated with age and showed a moderate correlation with IQ, suggesting that only the abilities captured by this subtest are more likely to be related to ones' overall level of intellectual functioning than age. The contribution of IQ, gender and years of education to performance on the TLC-E was not substantial to warrant even basic subdivision of norms according to those variables.

The norms provided here are within-group standard scores for the age groups: 20-34 years, 65-74 years and 75-89 years. These provisional norms were developed in accordance with the guidelines provided in the TLC-E manual. The manual advises that firstly, the normalized standard scores (z scores) should be derived from frequency distribution at each age interval. This should be accomplished by obtaining normalized percentile scores and converting them into z scores from the normal curve table. Each resultant z score should then be converted into a standard score using the formula (z x 3)+10= standard score. This would result in a normal distribution of standard scores with a mean of 10 and a standard deviation of 3 for each subtest (Wiig & Secord, 1989, p.48). In the present study scatterplots were first produced for each subtest for each age group to check for outliers. To examine the normality of the distribution of scores histograms were produced for each TLC-E measure in each age group. The examination of histogram plots confirmed the need to normalize the data. This was achieved by utilizing the Normal Distribution Fitting function in Statistica 6. The resultant expected percentile scores (for each age group) were converted into z scores from the normal curve table.

Table 10. Correlations between TLC-E performance and age, TLC-E performance and years of education and TLC-E performance and IQ.

TLC-E	Age	Education	WASI (2 subtest IQ)
Composite	-.61*	.35*	.29°
Ambiguous Sentences	-.59*	.32*	.27°
Making Inferences	-.61*	.29*	.32*
Recreating Sentences	-.73*	.44*	.26°
Figurative Language	-.24♦	.27°	.40*

Note: * $p<.02$, °$p<0.05$, ♦$p>0.05$

Each resultant z score was then converted into the standard score using the formula described above and a distribution of standard scores with a mean of 10 and a standard deviation of 3 was obtained (see Tables 11b-11d for scaled scores, TLC-E norms for the 18- year-olds are also provided for comparison in Table 11a). The scaled scores obtained for the reference group were generally comparable to those provided in the manual for the 18-year-olds, except that poorer performance was reflected by higher scores in the 20-34 year old reference group. The observed differences young-old and old-old adults' raw scores relative to the reference group is reflected in the distribution of the scaled scores shown in Tables 11 c and 11d. Relative to the 18- year-old norms, the most marked differences in the distribution of scores in the two elderly groups occurred for the Making Inferences and Recreating Sentences subtests. There was about one scaled score difference in the distribution on the Ambiguous Sentences, and as expected the distribution of scores on the Figurative Language was similar for the young and the elderly groups. There was no need to provide scaled scores for the TLC-E composite, as the composite represents a sum of scaled scores obtained on each of the TLC-E subtests. A standard composite score with a mean of 100 and a standard deviation of 15 can be obtained by first calculating the sum of the age-appropriate standard scores on all of the subtests and then using the TLC-E manual conversion chart to transform the resulting score.

3.2.4 *Internal consistency reliability of the TLC-E*

Internal consistency reliability coefficients are used to describe the precision of scores on the test. The internal consistency of the TLC-E was evaluated using the present sample data in order to establish whether any of the TLC-E items were not homogenous. Cronbach alpha calculated for the TLC-E measures indicated the presence of sound internal consistency for the Ambiguous Sentences, (0.78) Making Inferences (0.72) and Recreating Sentences (0.84) subtests. However the internal consistency coefficient for the Figurative Language subtest was low (0.41), which undermines the usefulness of this measure. The reduced reliability of Figurative Language is also likely to negatively affect the magnitude of correlations obtained between this subtest and other measures.

Table 11a. Norms provided by the TLC-E manual for ages 17-0 through 18-11

Subtest Standard Scores by Raw Scores Ages 17-0 through 18-11				
Subtest Standard Scores	**AS**	**MI**	**RC**	**FL**
17	--	--	--	--
16	--	--	--	--
15	39	--	78	--
14	--	--	77	36
13	38	36	--	35
12	37	35	75-76	34
11	35-36	34	74	33
10	33-34	33	72-73	32
9	30-32	32	71	30-31
8	29	31	68-70	28-29
7	26-28	30	65-67	25-27
6	24-25	28-29	61-64	22-24
5	22-23	26-27	58-60	21
4	19-21	25	54-57	16-20
3	1-18	1-24	1-53	1-15

Note: AS= Ambiguous Sentences, MI= Making Inferences, RC= Recreating Sentences, FL= Figurative Language

Table 11b. Provisional norms for the TLC-E for the reference (20-34 years) group.

Subtest Standard Scores by Raw Scores Ages 20-0 through 34-0				
Subtest Standard Scores	**AS**	**MI**	**RC**	**FL**
17	--	--	--	--
16	--	--	--	--
15	--	36	--	--
14	39	35	78	36
13	38	34	76-77	35
12	37	33	75	34
11	36	32	73-74	33
10	--	31	72	31-32
9	35	--	70-71	30
8	34	30	69	29
7	33	29	68	28
6	32	28	66-67	26-27
5	31	27	65	25
4	30	26	63-64	24
3	1-29	1-25	1-62	1-23

Note: AS= Ambiguous Sentences, MI= Making Inferences, RC= Recreating Sentences, FL= Figurative Language

Table 11c. Provisional norms for the TLC-E for the 65-74 years age group

Subtest Standard Scores by Raw Scores Ages 65-0 through 74-0				
Subtest Standard Scores	**AS**	**MI**	**RC**	**FL**
17	--	36	74-78	--
16	39	35	72-73	--
15	38	33-34	70-71	--
14	37	32	68-69	36
13	35-36	31	66-67	35
12	34	30	64-65	34
11	32-33	29	62-63	33
10	31	28	60-61	31-32
9	30	26-27	59	30
8	28-29	25	57-58	29
7	27	24	55-56	28
6	26	23	53-54	27
5	24-25	22	51-52	26
4	23	20-21	49-50	24-25
3	1-22	1-19	1-48	1-23

Note: AS= Ambiguous Sentences, MI= Making Inferences, RC= Recreating Sentences, FL= Figurative Language

Table 11d. Provisional norms for the TLC-E for 75-89 years age group

Subtest Standard Scores by Raw Scores Ages 75-0 through 89-0				
Subtest Standard Scores	**AS**	**MI**	**RC**	**FL**
17	--	--	77-78	
16	--	35-36	74-76	
15	39	33-34	71-73	36
14	36-38	31-32	68-70	34-35
13	34-35	29-30	65-67	33
12	32-33	28	62-64	32
11	30-31	26-27	59-61	30-31
10	27-29	24-25	56-58	29
9	25-26	22-23	53-55	28
8	23-24	21	51-52	26-27
7	21-22	19-20	48-50	25
6	18-20	17-18	45-47	23-24
5	16-17	15-16	42-44	22
4	14-15	14	39-41	21
3	1-13	1-13	1-38	1-20

Note: AS= Ambiguous Sentences, MI= Making Inferences, RC= Recreating Sentences, FL= Figurative Language

3.3 Working Memory Performance

A number of working memory measures was administered to the participants in attempt to better capture the overall construct of working memory. The tasks included: Daneman and Carpenter reading span, semantic fluency, the WMS-III working memory subtests and Digit Span.

3.3.1 *Daneman and Carpenter reading span.*

The Daneman and Carpenter reading span test assesses working memory capacity from the computational and storage perspective. Three measures of the Daneman and Carpenter span were derived: total number of last sentence words recalled, the traditional span score (see method section for calculation details), and a total number of true/false responses to the comprehension questions. Table 12 provides means and standard deviations for these measures. Significant group effects were detected for the total number of words recalled ($F (2, 59)= 28.8$ $p< 0.001$). The young adults recalled significantly more last sentence words than both of the elderly groups, but the two older groups did not differ from each other. The same pattern of results was obtained with a more traditional span score ($F (2, 59)= 5.78$, $p<0.05$). Again, the young adults significantly differed on their span score from the elderly groups who did not differ from each other. With regard to correct (true/false) responses it is important to note that most of the participants in each group obtained near perfect scores and there were no group differences ($F (2, 59)= 0.82$, n.s.) suggesting excellent comprehension of sentences by all groups.

Significant negative correlations were detected between the span measures and age (see Table 13), which is consistent with previously reported findings of correlations ranging from -0.4 to-0.7 (traditional span) (Carpenter, Miyake & Just, 1994), lending support to Just and Carpenter's theory (1980) that age is implicated in the decline of working memory. As the traditional span scores have reduced variability in comparison to the total word scores smaller correlations for the traditional span scores were obtained.

Table 12. Group means (standard deviations) on the Daneman and Carpenter reading span test.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F (2, 59)	p
Total words (max. 60)	35.85 (6.26)	25.45 (5.39)	23.80 (4.64)	28.8	0.001
Span (max 6)	2.88 (1.02)	2.23 (0.57)	2.18 (0.49)	5.78	0.04
Correct (T/F) (max.. 60)	58.30 (2.11)	58.95 (1.17)	58.45 (1.88)	0.82	0.44

Note: *F* ratios and *p* values are for one-way ANOVA

Table 13. Correlations between Daneman and Carpenter reading span task scores and age

	Total words	Span	True/false
Age	-0.73 *	-0.42*	0.65♦

Note: * p<0.02, ♦ p>0.05

Table 14. Correlations between TLC-E measures and Daneman and Carpenter total word and span scores

TLC-E	Total words	Span
Composite	.62	.47
Ambiguities	.62	.45
Inferences	.54	.39
Recreating Sent.	.69	.51
Figurative	.39	.37

Note: all *ps* <0.02

High correlations between the span measures and the TLC-E measures were also obtained (see Table 14) indicating that individuals with higher reading spans were better at the TLC-E tasks. The presence of a strong relationship was particularly evident for the Recreating Sentences subtest, which is not surprising given that this subtest requires the most organization and manipulation of material in working memory to produce a response. The Making Inferences and Ambiguous Sentences also demonstrated strong correlations with the span measures, whereas the Figurative Language was only moderately related.

High correlations between the reading span scores and TLC-E are important because they suggest that working memory capacity may contribute substantially to individual differences on this higher language task.

3.3.2 *WMS-III Working Memory Measures.*

Table 15 provides group means and standard deviations for the raw scores on Letter-Number Sequencing, Spatial Span and Digit Span subtests of the WMS-III.

Progressive age-related deterioration was observed on the Letter-Number sequencing subtest with significant pair-wise group differences also present ($F(2, 59) = 9.60$, $p<0.001$). Significant group effects were also found for the Spatial Span Total ($F(2, 59) = 17.5$, $p<0.001$), where the younger adults performed better than both of the elderly groups which did not differ from each other. The Spatial Span Total score consists of two tasks, Spatial Span Backwards and Spatial Span Forwards. A significant group effect was found for both Span Forwards ($F(2,59)=9.88$, $p<0.001$) and Span Backwards ($F(2, 59) = 14.51$, $p<0.001$). In both cases the young adults performed significantly better that the elderly groups which still did not differ from each other.

Whilst significant group effects were detected for the Digit Span Total score ($F(2, 59) = 6.30$, $p< 0.05$), only the group difference between the young and the old-old adults was significant. The Digit Span Total score is comprised of Span Backwards and Span Forward scores. Significant group effects were detected on Digits Forwards ($F(2, 59) = 9.60$, $p<0.01$), with pair-wise comparisons detecting differences between all three groups, but not on Digits Backwards ($F(2, 59) = 1.20$, *n.s.*).

Table 15. Group means (standard deviations) on the raw scores of the WMS-III working memory measures and the Digit Span.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F(2, 59)	p
L/N sequencing	12.45 (2.14)	10.27 (3.03)	8.45 (3.36)	9.60	0.001
Spatial Span total	18.50 (2.82)	14.55 (2.56)	13.55 (3.05)	17.50	0.001
Spatial Forwards	9.40 (1.88)	7.72 (1.61)	7.15 (1.50)	9.88	0.001
Spatial Backwards	9.10 (1.55)	6.82 (1.60)	6.15 (2.25)	14.51	0.001
Digit Span total	20.10 (2.67)	18.00 (3.99)	16.05 (3.97)	6.30	0.01
Digits Forwards	12.05 (1.88)	10.73 (2.21)	9.20 (2.04)	9.60	0.01
Digits Backwards	8.05 (1.96)	7.27 (2.6)	6.85 (2.74)	1.20	0.30

Note: *F* ratios and *p* values are for one way ANOVA

To summarize, the WMS-III working memory measures, only the Letter-Number Sequencing subtest was sensitive to age related progressive deterioration. A sharp decline was evident in the abilities measured by the Spatial Span from young to young-old which did not continue with further age. The performance of the three groups on Digit Span indicates that, contrary to the notion that Digits Backwards represents a more sensitive measure of working memory, it was not as sensitive to age differences as the Digits Forwards component.

The Wechsler Memory Scale –III provides tables for the conversion of raw scores of participants to the appropriate age scaled scores (see Table 16 for scaled score means; note that the manual does not provide separate scaled scores for Digits Backwards and Forwards). The WMS-III also combines the scores on Letter Number Sequencing and Spatial Span to arrive at the Working Memory Index (WMI) score (see Table 16). Given that the WMS-III was normed on the American population it was of interest to see if comparable results can be obtained with the New Zealand sample. The standard scores as well as the WMI index scores of the present sample were higher than those provided in the WMS-III (M=10, SD=3 for the subtest scores and M=100, SD=15 for the index score). No significant differences existed between the old-old elderly scaled scores and the population mean on Letter Number Sequencing, Digit Span and Spatial Span tests. The young-old and the young adults however scored significantly higher than the population mean ($p<0.05$) on all of these subtests. Although the differences between the sample and the population means generally fell at about 0.5 SD they are of potential clinical significance and indicate the presence of important differences in the New Zealand population relative to the American norms. Further investigations to confirm whether the current data reflects local differences are needed.

Apart from Digits Backwards, moderately high negative correlations (-0.42 to -0.62) were obtained between the raw WMS-III working memory scores and age (see Table 17), with Spatial Span total showing the strongest correlation.

Moderate to high significant correlations were obtained between the raw WMS-III working memory scores and the TLC-E subtests.

Table 16. Group means and (standard deviations) for working memory, Digit Span scaled scores and Working Memory Index on the WMS-III

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)
L/N	11.75 (2.53)	11.86 (3.06)	11.45 (3.43)
Spatial Total	11.35 (2.18)	11.45 (2.67)	11.30 (3.19)
Spatial Forwards	10.70 (2.64)	11.14 (2.64)	10.80 (2.83)
Spatial Backwards	11.90 (2.29)	11.54 (2.79)	11.35 (3.75)
Digit Total	11.75 (1.86)	12.13 (3.03)	10.90 (3.14)
WMI	108.85 (11.99)	109.95 (14.02)	108.65 (18.29)

Note: L/N= Letter/Number Sequencing, WMI= Working Memory Index

Table 17. Correlations between WMS-III working memory subtests and age

	Age
L/N Sequencing	-.51
Spatial Span total	-.62
Spatial Forwards	-.51
Spatial Backwards	-.58
Digit Span total	-.42
Digits Forwards	-.49
Digits Backwards	-.21

Note: all *ps* <0.02

Table 18. Correlations between the WMS-III working memory measures and digit span and TLC-E measures

TLC-E	DS	DF	DB	L/N	SS	SSF	SSB
Composite	.55	.53	.36	.61	.48	.34	.50
Ambiguities	.44	.42	.53	.57	.49	.40	.42
Inferences	.53	.55	.33	.48	.49	.42	.53
Recreating	.63	.58	.45	.63	.56	.37	.57
Figurative	.46	.45	.30	.39	.35	.31	.35

Note all *ps* <0.02, DS= Digit Span Total, DF= Digit Span Forwards, DB= Digit Span Backwards, L/N= Letter/Number Sequencing, SS= Spatial Span, SSF= Spatial Span Forwards, SSB= Spatial Span Backwards

The Recreating Sentences subtest showed the strongest associations with the WMS-III measures, while the Figurative Language obtained the weakest correlations (see Table 18). The pattern of correlations in general indicates that higher performance on the WMS-III working memory measures is associated with better performance on the TLC-E.

In summary, the results indicate that age is associated with reduced scores on the WMS-III measures of working memory. The Letter/Number Sequencing subtest appears to be particularly sensitive to deterioration with progressive age. The validity of the Digit Span subtest as the measure of general working memory (rather than a phonological subsystem) remains questionable. All of the WMS-III working memory measures correlated well with the TLC-E measure suggesting a working memory contribution to age differences on this discourse test.

3.3.3 *Semantic Fluency.*

The semantic fluency task purports to measure the controlled attention limits of the working memory system by assessing sustained effortful retrieval and the ability to shift sets. A number of measures were generated for the semantic fluency task: total number of words and switches and mean cluster size produced over a 5 minute period; number of words, switches and cluster size for each individual minute in task; cumulative increase in words produced, and the rate of change in word production (see Tables 19-21 for group means). To ascertain whether the groups differed in number of words produced over time a 3 (groups) x 5 (minutes in task) repeated measures ANOVA was conducted which produced a significant main effect for group ($F(2, 59)= 9.41$, $p<0.001$), a significant main effect for minutes in task ($F(4, 236)= 168.69$, $p<0.001$) and a significant group by minute interaction ($F(8, 236)= 12.65$, $p<0.001$). As evident from Table 19 significant differences between groups existed in the 1st, 2nd, and 4th minutes with young adults producing more words than the two elderly groups.

To better demonstrate group differences in word production over time and to enable the comparison with the presentation provided by Engle (1999) the words produced by each group were cumulated across minutes.

Table 19. Group means (standard deviations) for number of words produced in each minute on semantic fluency task

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	*p*
1st min.	25.55 (5.50)	19.73 (4.45)	18.10 (5.34)	0.001
2nd min.	14.05 (4.24)	11.64 (4.68)	9.10 (5.24)	0.001
3rd min.	9.35 (4.73)	9.91 (3.42)	8.15 (3.96)	0.36
4th min.	10.10 (5.01)	6.95 (3.76)	5.70 (2.52)	0.001
5th min.	6.70 (3.27)	7.64 (3.03)	5.20 (3.69)	0.66

Note *p* values are for simple main effects following a 3(group) x 5 (minutes in task) repeated measures ANOVA.

Table 20. Group means (standard deviations) for number of switches on semantic fluency task across time.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	*p*
1st min.	12.50 (4.22)	8.68(3.01)	7.69 (2.50)	0.001
2nd min.	5.65 (3.01)	3.58 (1.81)	3.30 (2.75)	0.001
3rd min.	4.15 (2.28)	3.45 (1.47)	2.95 (2.14)	0.61
4th min.	3.90 (2.67)	2.77 (2.22)	2.15 (1.72)	0.51
5th min.	2.40 (1.54)	2.09 (1.85)	1.50 (1.31)	0.21
Total	31.50 (7.96)	22.77 (6.38)	18.25 (6.98)	--

Note: *p* values are for simple mean effects following a 3(group) x 5 (minutes in task) repeated measures ANOVA.

Table 21. Group means (standard deviations) for mean cluster size on semantic fluency task across time.

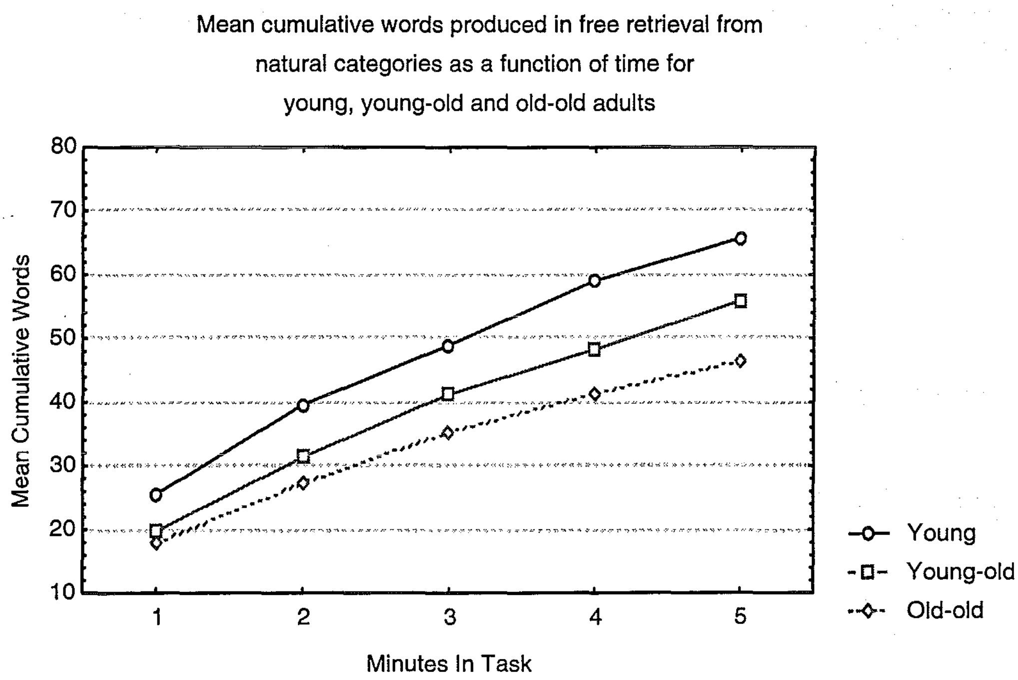
	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)
1st min.	1.23 (0.83)	1.38 (0.91)	1.72 (1.53)
2nd min.	1.54 (0.84)	1.85 (1.29)	1.48 (1.25)
3rd min.	1.18 (1.10)	1.65 (1.34)	1.52 (1.93)
4th min.	1.55 (1.37)	1.10 (0.90)	1.62 (1.36)
5th min.	1.44 (1.85)	2.87 (3.28)	1.50 (1.53)
Total	1.33 (0.51)	2.55 (0.67)	1.19 (2.85)

Note: all *p* values for simple main effects following a 3 (group) x 5(minutes in task) repeated measures ANOVA were>0.05

Figure 1 demonstrates the differences in word recall between groups over time. In the first minute the young adults recalled more words than the two elderly groups which did not differ from each other. As time progressed the level of disparity between the groups appeared to increase over the 5 minute period. By the fifth minute the young-old adults had recalled significantly fewer words than the young adults overall, whilst the old-old adults not only significantly differed in their recall from young adults but also from the young-old. In his study Engle (1996) investigated group differences in word production by diving the sample into the high and low span groups based on the participants' performance on the operations span. The participants in the present study were also divided into high (upper quartile) or low (lower quartile) working memory capacity groups based on the Daneman and Carpenter total word score. In the present study the range for the lower quartile was 14-23, M =19.42, SD= 2.31and for the upper quartile, 33-50, M=38, SD=4.56. When the semantic fluency performance of the span groups was analyzed a significant main effect for group ($F(1, 29)=9.19$, $p<0.001$), significant main effect for minutes ($F(4, 116)=84.89$, $p<0.001$) and a significant group by minute interaction ($F(4, 116)= 5.96$, $p<0.001$) was detected. Figure 2 demonstrates that in the first minute the differences between the low and high span participants were small but as the time progressed the level of disparity between the groups has increased. The results obtained in the present experiment compare well with those generated by Engle (1999), who also found that differences between low and high span participants increased by the 5^{th}-6^{th} minute of the task.

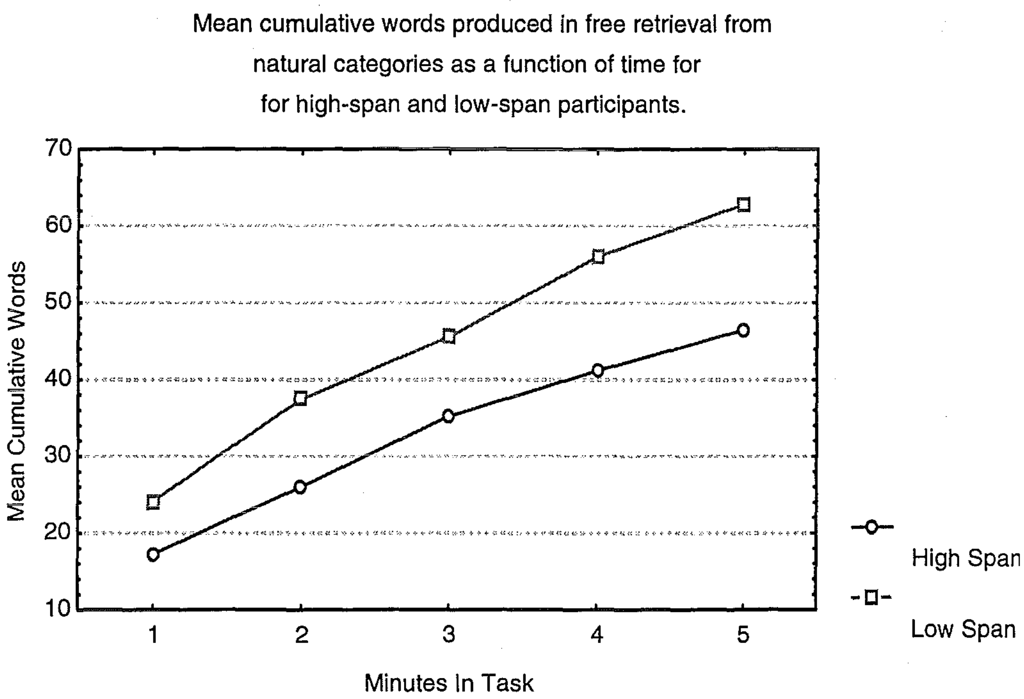
One of the assumptions made by Engle's General Capacity theory is that the high and low span participants differ in the amount of activation available to the working memory system. The theory predicts that the low span individuals would experience more difficulty spreading the activation over time, and hence would demonstrate substantial reductions in word production rate in comparison the high span participants. The present study examined group differences in production rate. To obtain a measure of change in production rate, the number of words produced at each minute past the first minute was converted into a percentage of words relative to the first minute.

Figure 1.



Mean cumulative words produced in free retrieval from
natural categories as a function of time for
young, young-old and old-old adults
Mean Cumulative Words
80
70
60
50
40
30
20
10
1
2
3
4
5
Minutes In Task
Young
Young-old
Old-old

Figure 2.

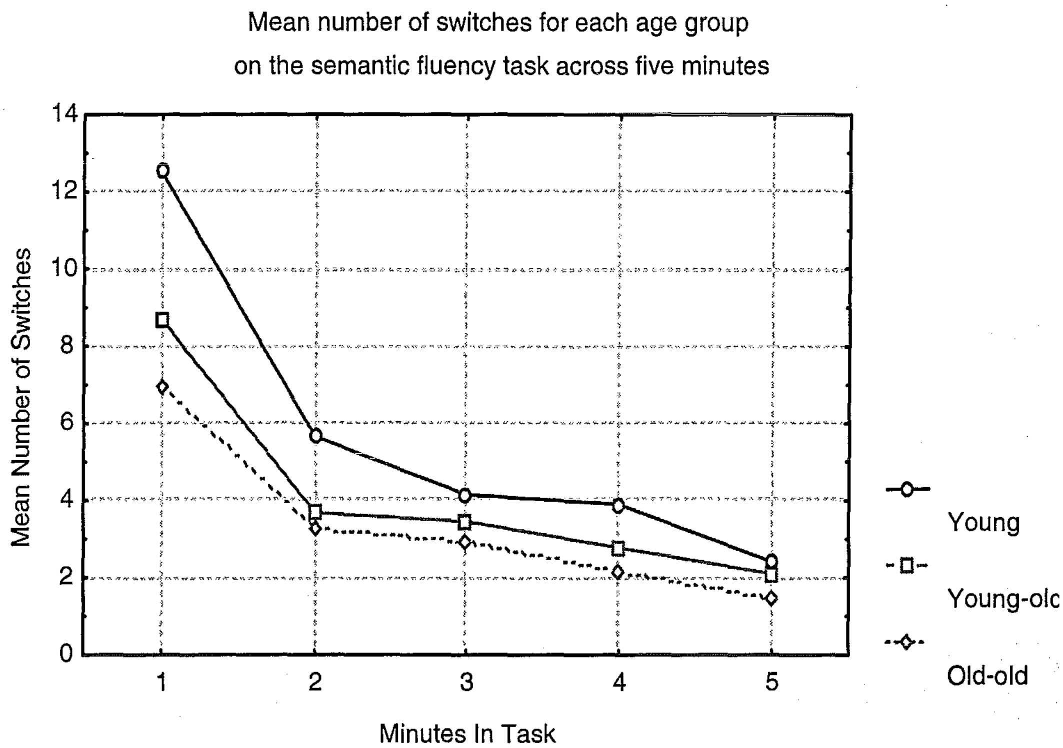

Mean cumulative words produced in free retrieval from
natural categories as a function of time for
for high-span and low-span participants.
70
60
50
40
30
20
10
Mean Cumulative Words
1
2
3
4
5
Minutes In Task
High Span
Low Span

A total rate of change measure represented an average percent decrement in word production (in comparison to the first minute) over the last four minutes. A repeated measures 3(group) x 4(minutes in task) ANOVA revealed an absence of significant main group effect ($F(2, 59)=1.81$, n.s.) or group by minute interaction ($F(6, 117)=2.09$, n.s.). However a significant main effect for minutes in task was obtained ($F(3, 117)= 24.37$, $p<0.01$). A simple main effects analysis indicated an absence of significant group differences in the rate of change for minutes 2, 3, and 4. In the fifth minute the young-old group demonstrated a significantly higher rate of change than the young or the old-old. These data suggest that the rate of change in word production across minutes does not substantially differ between groups. Hence the differentiation observed between three groups in the fifth minute, when cumulative scores are examined, is a function of an overall reduced number of words generated by the elderly groups rather than a differential rate of change in production.

Troyer et al (1997) have recently evaluated the contribution of switching and clustering to the performance on semantic fluency task, arguing that both components are essential for successful accomplishment of the task. Troyer et al. (1997) utilized a 1-minute duration task. To allow comparison with Troyer's data, clustering and switching was examined in the first minute as well as in the whole 5 minutes. Group means and standard deviations for the number of switches and for cluster size are shown in Tables 20 and 21 respectively. Significant group effects were found for the number of switches made in the 1st minute ($F(2, 59)= 8.36$, $p<0.05$). The younger adults switched more often than the elderly groups who did not significantly differ from each other. The smaller number of switches produced by the elderly could be a reflection of a smaller number of words generated on the task. However there was no significant group effect for the mean cluster size for the first minute detected ($F(2, 59)<1.0$, *n.s.*). Thus, in the first minute, age was differentially associated with the two components. The age differences in switching favored the young, whereas there were no age differences in clustering. These results are consistent with what has previously been reported by Troyer et al. (1997).

Figure 3.

Figure 3.



When the performance of the groups on switching and clustering was examined over the whole 5-minute period the same findings were observed. A 3(group) x 5 (minutes in task) repeated measures ANOVA on switching detected significant main effects for groups ($F (2, 59)= 15, 81, p<0.001$), and minutes ($F(4, 236)= 107.59, p<0.001$) and a significant group by minute interaction ($F(8, 236)= 4.02, p<0.001$). As Figure 3 indicates the young adults produced a higher number of switches than both groups in the first and second minute but as the time progressed switching decreased with differences on this measure becoming marginal as testing progressed.

The differences in cluster size produced by the three groups was evaluated by a 3 (groups) x 5 (minutes in task) repeated measures ANOVA, which produced a non-significant main effect for groups ($F (2, 59)=1.14$, *n.s.*), minutes ($F(4, 236)= 1.40$, *n.s.*) and group by minute interaction ($F(8, 236)=1.74$, *n.s.*), suggesting that there were no differences between groups in mean cluster size in the test.

The data on switching and clustering was re-examined using the span score (low/high) quartiles as a grouping variable. A 2 (groups) x 5 (minutes in task) repeated measures ANOVA on clustering scores detected an absence of significant main effect for group ($F(1, 29)= .20$, n.s.), minutes ($F(4, 116)= .44$, n.s.) or group by minute interaction ($F(4, 116)= .66$, n.s.). The analysis of switching data demonstrated evidence of significant main effect for high/low span group ($F(1, 29)= 10.32, p<0.001$), and minutes ($F(4, 116)= 50.73, p<0.001$) but not for group by minute interaction($F(4, 116)= 1.56$, n.s.).

Intercorrelations between fluency measures are presented in Table 22. The rate of change correlated significantly only with the number of words produced. A significant, moderately high correlation was detected between the number of words produced over 5 minutes and the number of switches produced over 5 minutes. There was an absence of association between cluster size over 5 minutes and the total number of words generated, contrary to Troyer et al.'s (1997) report of moderate (.37) but significant correlations between cluster size and total words produced. The correlations were re-examined for just the first minute in task, and same pattern was obtained: switching was correlated significantly ($p<0.02$) with number of words (0.54), while clustering was not (0.20, $p>0.05$).

Table 22. Intercorrelations between fluency measures and age

	Age	Words 5min	Clusters 5min	Switches 5min	Rate of change
Age	--				
Words 5 min	-.47*	---			
Cluster 5 min	.19♦	.05♦	--		
Switches 5 min	-.60*	.63*	-.43*	--	
Rate of change	.03♦	.30*	.04♦	.15♦	--

Note * p<0.02, ♦p >0.05

Table 23. Correlations between Semantic Fluency measures and the TLC-E measures.

TLC-E	Fluency Words	Switching	Rate of Change	Clustering
Composite	.44*	.43*	.05♦	-.05♦
Ambiguities	.50*	.47*	.13♦	-.09♦
Inferences	.43*	.43*	-.07♦	-.30°
Recreating sent	.53*	.53*	.04♦	-.16♦
Figurative	.29°	.19♦	.16♦	-.16♦

Note: * p<0.02, ° p<0.05, ♦ p>0.05.

The obtained pattern of correlations suggests that switching is more important than clustering for optimal performance on the fluency task. Age demonstrated strong and significant correlations with switching and a total number of words produced (see Table 22). None of the other measures were correlated with age significantly. All of the TLC-E measures apart from Figurative Language were moderately highly correlated with switching and the total number of words produced (see Table 23).
Figurative Language was not significantly related to the switching measure and was only mildly related to the number of words produced. The Making Inferences subtest was the only TLC-E subtest that was significantly correlated with clustering. The rate of change on the fluency task was not significantly correlated with any of the TLC-E measures.

To establish which of the fluency measures best represent participants' performance on the semantic fluency task a reliability analysis was conducted. The reliability of the composite that included total words generated, total switches, total clusters and the rate of change measure was $\alpha = 0.51$. However, a substantial improvement in reliability ($\alpha = 0.70$) was attained by deletion of the rate of change and cluster size measures from the composite. Governed by these reliability estimates as well as the observation that switching and total words demonstrated superior sensitivity to age related decline, switching and total words over the 5 minutes test were selected for the composite to represent performance on the fluency task.

In summary, the performance of age groups on the semantic fluency task suggested that elderly individuals experience decline in controlled effortful retrieval of semantic information including the ability to switch successfully between retrieval strategies. This decline in controlled sustained attention functions was in turn associated with the deterioration of performance on three out four TLC-E subtests (excluding Figurative Language).

3.3.4 *Interrelationship between working memory mea*sures.

Correlations between the different working memory measures indicated the presence of significant moderate to high interrelationships (see Table 24). The substantial degree of overlap between the measures suggests that they all possibly assess the same construct, and indicates that the measures can be successfully combined into a single composite.

Table 24. Intercorrelations between Working Memory measures

	D&C	Digit	Fluency W	Switches	L/N	Spatial Span
D& C	--					
Digit	.61	--				
Fluency W	.40	.49	--			
Switches	.41	.41	.63	--		
L/N	.56	.57	.50	.37	--	
Spatial Span	.54	.40	.44	.45	.52	--

Note all *ps* <0.02, Fluency W= total number of words produced in the fluency task over 5 minutes, L/N=Letter/Number Sequencing, D&C= Daneman and Carpenter total word Span, Digit=Digit Span total score

A single working memory composite was thought to better represent working memory than each individual measure in terms of being more effective when the examinations of the relationship of working memory to the TLC-E were conducted (and see section 3.7).

3.4 Processing Speed Performance

Speed of processing is thought to be one variable that possibly mediates age differences on discourse tasks. Three measures of speed of processing were used in the present study, word reading, color naming and Digit Symbol Coding. Participants' mean performance on the processing speed measures is presented in Table 25. The younger adults performed significantly better ($F(2, 59) = 40.95$, $p<0.001$) on the Digit Symbol task than both elderly groups, which also significantly differed from each other, suggesting the presence of progressive deterioration on this speed measure with age. The raw scores on the Digit Symbol task were also converted into the age-appropriate scaled scores from the WAIS-III tables. No significant differences were detected when the obtained mean scaled scores for each age group were compared to their respective population mean ($M= 10$, $SD=3$).

A significant group effect was obtained for the number of black ink word read in the 45 seconds of the Stroop task ($F(2, 59)= 11.54$, $p<0.001$), this time with only the old-old group scoring significantly lower than both the young and the young-old. By contrast, on the congruent color naming task all the groups significantly differed from each other ($F(2, 59)= 27.13$, $p<0.001$) with the young adults obtaining the highest scores.

As would be expected all three speed measures were highly intercorrelated (see Table 26). Performance on the speed tasks also demonstrated strong negative associations with age, with correlations ranging between -0.43 to -0.77.

Digit Symbol and color naming showed high correlations with all of the TLC-E measures, apart from Figurative Language (see Table 27). Word reading was significantly but moderately related to the TLC-E. Thus, the participants who were better at speed of processing tasks were also more likely to score well on the TLC-E.

The processing speed measures evidenced moderate to high significant relationship with the working memory measures (see Table 28). Digit Symbol was particularly well correlated with the working memory tasks.

Table 25. Group means and (standard deviations) for Digit Symbol raw and scaled scores and Color naming and Word reading scores

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F (2, 59)	p
Digit Symbol					
Raw	81.45 (12.17)	56.73 (9.37)	45.50 (16.88)	40.95	0.001
Scaled	10.70 (2.23)	11.04 (1.96)	11.25 (2.79)	--	--
Word reading					
	101.5 (14.10)	97.45 (9.58)	81.75 (16.36)	11.54	0.001
Color naming					
	77.65 (12.61)	62.72 (11.50)	50.35 (11.02)	27.13	0.001

Note: *F* ratios and *p* values are for one-way ANOVA

Table 26. Correlations between speed of processing measures and age.

	Digit Symbol	Words	Colors	Age
Digit Symbol	--			
Word score	.67	--		
Color score	.76	.67	--	
Age	-.77	-.43	.70	--

Note: all *p*s <0.02

Table 27. Correlations between speed of processing measures and the TCL-E measures.

TLC-E Measures	Digit Symbol	Word	Color
Composite	.60	.44	.62
Ambiguities	.66	.48	.64
Inferences	.64	.46	.59
Recreating Sent.	.65	.44	.64
Figurative	.38	.33	.37

Note all *p*s are <0.02

Table 28. Correlations between measures of processing speed and working memory.

	Digit Symbol	Word	Color
D&C	.66	.42	.65
Digit Span	.60	.39	.52
Fluency words	.61	.59	.57
Switching	.57	.33	.49
L/N Sequencing	.61	.57	.66
Spatial Span	.64	.40	.53

Note all *ps* <0.02 D&C= Daneman and Carpenter reading span total word score, L/N Sequencing = Letter/Number Sequencing.

The pattern of correlations suggests that processing speed contributes to working memory performance.

In summary, age related declines were observed on all of the measures of processing speed, with progressive declines evident on more complex speed measures (e.g. Digit Symbol and color naming). Good performance on the speed tasks was found to be related to one's discourse skills as well as working memory function.

3.5 Inhibitory Efficiency Performance

Inhibitory efficiency was another factor examined in the present study as possible contributor to age differences on the TLC-E. The ability to inhibit irrelevant information is thought to play a part in efficient language processing. The Stroop color word task provided the measure of inhibitory efficiency. Participants' mean performance on the Stoop test is presented in Table 29. Group differences on the color and word scores control conditions of the Stroop are already described in the section on procession speed. Consistent with early research on the Stroop phenomenon, reading words was faster than naming colors or reading incongruent words (MacLeod, 1991). The young adults performed better than the elderly groups on the color-word condition ($F(2, 59)=44.19$, $p<0.001$) and the young-old adults also obtained higher scores than the old-old. The color-word score is thought to reflect the ability of the person to inhibit an inappropriate response. However, the age differences in the color-word score may simply be a reflection of differences in the speed of reading rather than in inhibitory efficiency. To control for the effects of the speed of reading analysis of covariance (ANCOVA) was conducted with the word and the color scores as covariates which showed that the significant group difference on the color-word score was reduced but maintained ($F(2, 57)= 13.33$, $p<0.01$). Another way to compare the groups on their inhibitory efficiency is to calculate an interference score. The formula for calculating the score (see in the method section) controls for the speed of responding and is justified given the within subject relative color and word scores. The interference scores obtained by the groups are provided in Table 29.

Table 29. Group means (standard deviations) for word, color, color-word and interference scores on STROOP.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F (2, 59)	p
Words	101.50(14.11)	97.45 (9.58)	81.75 (16.56)	11.54	0.001
Colors	77.65 (12.61)	62.72 (11.50)	50.35 (11.07)	27.13	0.001
Color-Words	48.35 (9.01)	31.23 (7.77)	24.10 (8.43)	44.19	0.001
Interference	6.30 (8.12)	-6.75 (5.75)	-6.68 (6.06)	23.76	0.001

Note: *F* ratios and *p* values are for one-way ANOVA

Table 30. Correlations between TLC-E measures and Stroop interference score.

TLC-E	Interference	Color-word
Composite	.28°	.53*
Ambiguities	.34*	.59*
Recreating Sent.	.43*	.61*
Inferences	.46*	.64*
Figurative	.10♦	.30*

Note: *- $p<0.02$, °- $p<0.05$, ♦-$p>0.05$

A score below zero indicates poor resistance to interference. A significant group effect ($F(2, 59) = 23.76$, $p<0.001$) was detected on this interference score, with the young group obtaining higher interference scores than both of the elderly groups which did not differ from each other. Thus the older groups were equally far less efficient at inhibiting inappropriate responses than the young adults. As expected the color-word score and the interference score were highly correlated with age (–0.80 and –0.67 respectively).

Color-word reading scores demonstrated strong associations with the TLC-E measures apart from Figurative Language (see Table 30). By contrast the interference scores showed weaker associations with the TLC-E measures.

In summary, the results indicated that the ability to inhibit irrelevant responses declines sharply from young to young-old, but shows virtually no deterioration with further age. As the color-word score contains an element of speed it is likely that it was the speed of responding that resulted in higher correlations between the color-words and the TLC-E. When the speed of responding was controlled for (as in the interference score) the relationship with the TLC-E was less pronounced suggesting that inhibitory efficiency has only limited associations with discourse skills. That is, inhibitory efficiency as assessed by the Stroop task seems likely to contribute only to a limited degree to the performance on the TLC-E.

3.6 Delayed and Immediate Auditory Memory Performance

Long-term memory may also play a part in mediating the performance on discourse tasks, particularly when retrieval of knowledge is required for successful comprehension of material. In addition, according to Cantor and Engle (1993), long-term memory retrieval abilities may also be crucial for efficient working memory functioning. To evaluate the contributions of long-term memory the performance of younger and older adults was examined and compared on the WMS-III Auditory Delayed and Immediate memory subtests (Logical Memory I and II and Verbal Paired Associates I and II). Table 31 presents raw score means (standard deviations) obtained by different age groups on these subtests.

As evident from Table 31 significant group differences were detected on the Logical Memory I ($F (2, 59)= 9.98$, $p<0.001$) and Logical Memory II ($F (2, 59) = 16.29$, $p<0.001$) subtests. On both immediate and delayed versions of the test the young adults performed significantly better then the elderly groups which did not differ from each other. Significant group effects were also observed for Verbal Pairs I ($F (2, 59) =16.21$, $p<0.001$) and Verbal Pairs II ($F (2, 59)= 27.39$, $p<0.001$). In contrast to the Logical memory test, both of these immediate and delayed versions of the paired associates test revealed that all three age groups significantly differed from each other.

The raw scores obtained by the groups on the Logical Memory and Verbal Pairs subtests were converted into the age appropriate scaled scores from the WMS-III tables. The scaled scores on the Logical Memory I and Verbal Pairs I were combined to produce an Immediate Auditory Memory Index. The Delayed Auditory Memory Index was derived by combining the scaled scores the Logical Memory II and Verbal Pairs II subtests. The mean scaled scores obtained by the three age groups (see Table 32) were compared to the WMS-III population means ($M =10$, $SD=3$ for subtest scores and $M=100$, $SD= 15$ for index scores). No significant differences between the WMS-III population means and the present sample means were detected on the Auditory Immediate Index for all age groups and on the Auditory Delayed Index for the old-old group. However the young-old and the young adults scored significantly higher ($p<0.05$) on the Auditory Delayed Index subtests than the WMS-III population, although the differences did not exceed 1 SD. The finding that 20-34 year old and 65-74 year old New Zealanders tend to perform slightly better on the Delayed Auditory Memory subtests is of potential clinical importance and further highlights the need for the development of New Zealand WMS-III norms.

Not surprisingly all four measures of Auditory Memory were strongly intercorrelated (see Table 33). The Auditory Memory subtests also demonstrated moderately high correlations with age (see Table 33).

Immediate and delayed memory measures demonstrated reasonably strong correlations with discourse tasks. As shown in Table 34 the correlations for the TLC-E measures ranged between 0.36 to 0.58.

Table 31. Group means (standard deviations) for raw score on the Auditory Immediate and Auditory Delayed subtests of the WMS-III

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)	F (2,59)	p
LMI	44.70 (9.41)	33.77 (10.35)	31.05 (10.96)	9.98	0.001
LMII	29.25 (7.40)	19.27 (6.34)	17.75 (7.14)	16.29	0.001
VPAI	22.20 (7.98)	15.68 (7.94)	9.35 (4.99)	16.21	0.001
VPAII	7.20 (1.47)	5.50 (2.11)	3.00 (1.69)	27.39	0.001

Note: LMI and LMII= Logical Memory subtests I and II, VPAI and VPAII= Verbal Paired Associates subtests I and II. *F* ratios and *p* values are for one-way ANOVA.

Table 32. Group means (standard deviations) for scaled scores on the Auditory Delayed and Auditory Immediate subtests and Index scores on the WMS-III.

	Young (20-34y)	Young-old (65-74y)	Old-old (75-89y)
LMI	10.90 (2.51)	9.82 (2.95)	10.40 (3.22)
LMII	11.85 (2.41)	10.23 (2.59)	11.30 (2.51)
VPAI I	11.20(3.30)	11.14 (3.41)	9.05 (1.99)
VPAII	11.70 (2.00)	12.09 (2.58)	9.80 (1.96)
AI	105.70 (14.11)	102.59(15.35)	95.95 (16.78)
AD	109.95 (11.22)	106.50(13.08)	102.90(10.63)

Note: LMI and LMII= Logical Memory subtests I and II, VPAI and VPAII= Verbal Paired Associates subtests I and II, AI= Auditory Immediate Index, AD= Auditory Delayed Index.

Table 33. Intercorrelations between the WMS-III Auditory Memory Subtests

	LMI	VPAI	LMII	VPAII	Age
LMI	--				
VPAI	.58	--			
LMII	.85	.68	--		
VPAII	.53	.86	.66	--	
Age	-.52	-.57	-.61	-.64	--

Note all *p*s<0.02, LMI and LMII= Logical Memory subtests I and II, VPAI and VPAII= Verbal Paired Associates subtests I and II.

Table 34. Correlations between delayed memory scores and TLC-E measures

TLC-E Measures	LMI	VPAI	LMII	VPAII
1. Composite	.51	.46	.52	.54
2. Ambiguities	.52	.54	.54	.58
3. Inferences	.56	.48	.45	.51
4. Recreating Sent.	.48	.49	.53	.53
5. Figurative	.45	.36	.47	.40

Note all *ps* <0.02 LMI and LMII= Logical Memory I and II, VPAI and VPAII= Verbal Paired Associates I and II.

Table 35. Intercorrelations between Auditory Delayed memory subtests and working memory measures.

	LMII	VPAII
D &C	.61	.60
Digit Span	.50	.51
Fluency words	.50	.58
Switching	.41	.56
L/N Sequencing	.54	.57
Spatial Span	.49	.45

Note all *ps* <0.02, D& C= Daneman and Carpenter reading span total words score, L/N Sequencing Letter/Number Sequencing, LM2= Logical Memory II, VPAII= Verbal Paired Associates II.

Strong associations between the working memory measures and Auditory Delayed subtests were detected (Table 35). The Daneman and Carpenter total word span score and Letter/number sequencing in particular, demonstrated the strongest degrees of association with the Delayed Memory subtests. Notably Engel's (1999) measure of retrieval (total number of words produced on the semantic fluency task) showed good relationship with the Delayed Memory subtests, suggesting that the participants who were better at retrieving the exemplars of one category were also better at recalling newly learned information after delay.

In summary, the obtained results confirmed previous findings that older adults experience decline in LTM retention. The LTM performance was also associated with working memory efficiency as well as the ability to produce and comprehend language.

3.7 Contribution of mediating variables to age differences on the TLC-E

One of the goals of the present study was to establish to what degree mediating variables such as working memory, speed of processing, LTM and resistance to interference contribute to the performance on the TLC-E. For this purpose, firstly composite scores for each mediating variable were obtained. Composite scores should better reflect the relevant underlying theoretical constructs than any single test measure. To ascertain what measures should be combined into any given composite the correlations between each single measure were first examined. Then, following the method employed by Salthouse (1992), the raw scores on each selected subtest were converted into z scores, which were then averaged to arrive at a final composite. The reliability of this composite was then calculated; a reliability over 0.7 was considered acceptable.

Two main statistical techniques were employed to assess the contribution of each composite to the TLC-E performance, analysis of covariance and path analysis.

The working memory composite: Out of all the working memory measures obtained in the present study clustering and rate of change on the semantic fluency task were not included in the composite. These two measures demonstrated very poor associations with the other working memory scores (correlations ranging from 0.04 to 0.20). Additionally, examination of the reliability of the composite suggested that exclusion of these measures would substantially improve the reliability (increase in α from 0.66 to 0.74). Hence the measures that comprised the working memory composite included: Letter/Number sequencing, Spatial Span Total, Digit Span Total, Daneman and Carpenter total words, and semantic fluency total number of words and switches produced in 5 minutes. The correlations between selected measures were moderate to high ranging from 0.37 to 0.61 (see Table 19), suggesting that the measures can be effectively combined to represent a single construct. Examination of item total statistics suggested that deletion of any of the six measures would not result in substantial improvement in the alpha level.

The speed of processing composite: Digit Symbol Coding, word reading and color naming tasks were included in the speed of processing composite. The correlations between the tests were acceptable (see Table 31) suggesting that they can be grouped into a single composite. The reliability coefficient for the processing speed composite was highly robust α=0.87

The long-term memory composite: The raw scores obtained by the participants on the Logical Memory II and Verbal Pairs II constituted the long-term memory composite. The correlation between the two subtests was 0.66. The reliability estimate was robust (α=0.87).

Inhibition: Because the color-word score on the Stroop task was highly correlated with the speed measures, it seemed inappropriate to use it as a measure of inhibition. Hence, the Stroop interference raw score (i.e. not converted into z score) was the only measure of inhibition used.

The discourse composite: The TLC-E composite raw score was used as a composite measure of language/discourse functioning. Although the subsets that constitute the composite score measure different abilities, the reliability of the composite was acceptable α =0.79 and the examination of item total statistics suggested that deleting any of the four subtests would not result in substantial improvement in the reliability estimate.

3.7.1. *Analysis of covariane on the TLC-E*

Analysis of covariance (ANCOVA) adjusts the raw score means, based on covariate variables, before testing for significance between groups. Prior to examining the contribution of composite measures to the TLC-E performance, an ANCOVA for each composite with years of education was conducted. Years of education was the only main extraneous variable on which groups differed due to population characteristics. The ANCOVA indicated that Group remained an influential variable on working memory composite ($F(2, 58)=24.88$, $p<0.001$), speed composite ($F(2, 58)=24.30$, $p<0.001$), interference ($F(2, 58)=22.45$, $p<0.001$) and LTM composite ($F(2, 58)=18.94$, $p<0.001$).

Entering the working memory composite score as a covariate resulted in group effects reducing to non-significant levels on the TLC-E composite score ($F(2, 58)=1.95$, n.s.), as well as on the Ambiguous Sentences ($F(2, 58)=0.90$, n.s.) and Making Inferences ($F(2, 58)=2.32$, n.s.) subtests. The group effects on the Recreating Sentences subtest remained significant ($F(2, 58)=4.11$, $p<0.02$) though the *F* ratio was substantially reduced. Indeed, the group effect on Recreating Sentences was maintained ($F(2, 55)=4.30$, $p<0.02$) even when all four covariates (working memory, speed, interference and LTM) were entered into the ANCOVA.

Thus differences in working memory substantially accounted for group differences in language competence as reflected in TLC-E measures, where clear initial age differences were already apparent.

When the processing speed composite was entered as a covariate, group effects were maintained at a significant level for the TLC-E composite score ($F(2, 58)=2.24$, $p<0.05$) and for the Recreating Sentences subtest ($F(2, 58)=9.44$, $p<0.001$) though in both cases the *F* ratios were reduced. Group effects were no longer significant for the Ambiguous Sentences ($F(2, 58)=0.90$, n.s.) or Making Inferences ($F(2, 58)=2.32$, n.s.) subtests when differences in processing speed were accounted for.

Entering the Stroop interference score as a covariate by itself had relatively less effect on the observed group effects on any of the TLC-E measures: TLC-E composite ($F(2, 58)=14.94$, $p<0.0001$), Making Inferences ($F(2, 58)=8.72$, $p<0.0001$), Ambiguous Sentences ($F(2, 58)=9.41$, $p<0.0001$) and Recreating Sentences ($F(2, 58)=18.33$, $p<0.0001$).

The addition of the LTM memory composite as a covariate substantially affected group differences only on the Ambiguous Sentences subtests ($F(2, 58)=2.38$, n.s.) but with milder effects on other TLC-E measures: TLC-E composite ($F(2, 58)= 4.53$, $p<0.02$), Making Inferences ($F(2, 58)=5.28$, $p<0.01$), Recreating Sentences ($F(2, 58)=12.15$, $p<0.01$).

As the Figurative Language subtest did not produce significant group effects the influence of covariates was difficult to estimate. The addition of any one covariate resulted only in small attenuation of the *F* ratio and not more than 1 point increase in the means of the elderly and 1 point reduction in the means of the young.

In summary, the ANCOVA results support the notion of working memory contribution to group differences on the TLC-E. Working memory was the only covariate that successfully accounted for age differences on the TLC-E composite and had the greatest impact on age differences on the various subtests. The working memory rival, speed of processing, contributed to age related differences in Making Inferences and Ambiguous Sentences but was not sufficient to account for group differences on the TLC-E composite. The long-term memory mediated group differences on the Ambiguous Sentences, whilst resistance to interference had less influence on the observed group effects on the TLC-E composite or subtest scores. None of the covariates individually or in combination were able to successfully account for age differences on the Recreating Sentences subtest.

3.7.2 *Path analyses models*

<u>Path models for the TLC-E composite</u>

A path model depicts how a particular set of independent variables influence a dependent variable under consideration. The simple correlations between the measures used in the proposed path analyses are presented in Table 36. Age showed high zero-order correlations with the TLC-E composite and with all potential mediator variables (working memory composite, speed composite, LTM composite, and Stroop inhibition). Unlike the other potential mediator variables, Stroop Inhibition had only a weak association with TLC-E performance.

Table 36. Zero order correlations between measures used in path analysis

	1	2	3	4	5	6
1. Age	--					
2. TLC-E	-.61	--				
3. Inhibition	-.67	.28°	--			
4. Working memory	-.74	.69	.51	--		
5. Speed	-.71	.62	.45	.80	--	
6. Long-term memory	-.69	.58	.50	.77	.73	--

Note °$p<0.05$, all other *ps* <0.02

In the first set of path analyses, the influence of each potential mediator variable on the association between age and TLC-E composite was examined separately. The basic model (Figure 4) examines the relationship between age on language functioning when mediated by an intervening variable (X). Note that an asterisk in the path diagrams signifies $p<0.05$. Note that in all path diagrams, the standard error is presented in brackets adjacent to its path coefficient.

Figure 4. Path diagram illustrating hypothesized relationship among age, TLC-E and a mediator variable (X).

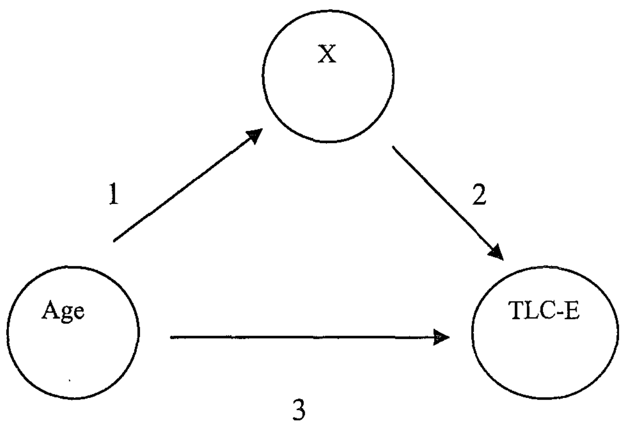


Figure 5. Path diagram illustrating relations among age, TLC-E and working memory.

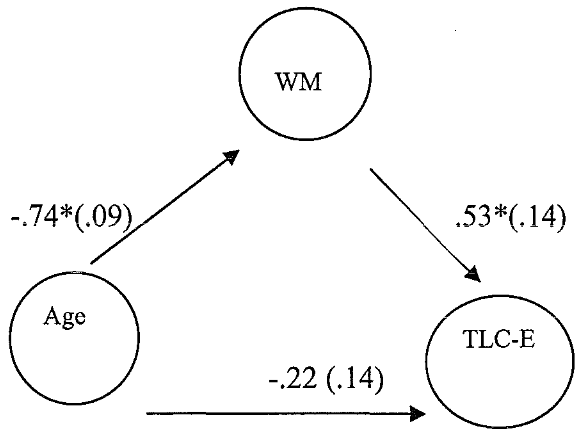


Figure 5 illustrates the influence of working memory as a sole mediator of the relationship between age and TLC-E performance. The strong association between working memory and TLC-E scores ($r = 0.69$, $p<0.001$) remained after controlling for the effect of age ($\beta = 0.53$, $p<0.001$). The association between age and TLC-E ($r = 0.61$, $p<0.001$) was substantially reduced to a non-significant level after controlling for age differences in working memory ($\beta = - 0.22$, n.s.).

Figure 6. Path diagram illustrating relations among age, TLC-E and speed of processing.

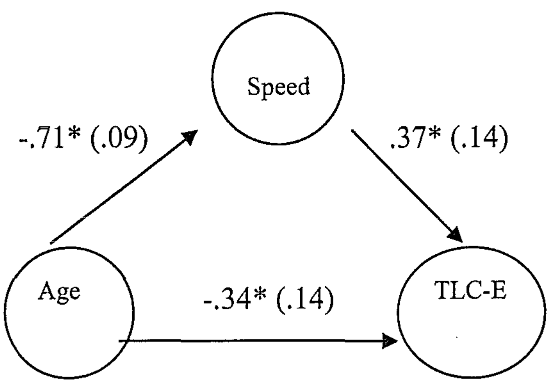
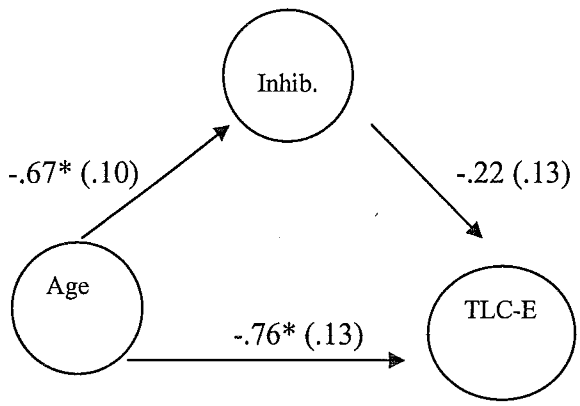


Figure 6 illustrates the influence of speed as a sole mediator of the relationship between age and TLC-E performance. The strong association between faster speed of processing and higher TLC-E scores (r = 0.62, p<0.001) remained after controlling for the effect of age (β = 0.37, p<0.01). The association between age and TLC-E (r = 0.61, p<0.001) was reduced after controlling for age differences in speed, but remained significant (β = - 0.34, p<0.05). That is, working memory was more successful than speed in accounting for age differences on the TLC-E (compare Figures 5 and 6).

Figure 7. Path diagram illustrating relations among age, TLC-E and inhibition.



Although inhibitory efficiency (Stroop interference score) showed a weak, positive zero-order association with TLC-E (r = 0.28, p<0.05), this relationship was reversed when the effects of age were partialled out (β = - 0.22 n.s.; Figure 7). Moreover, controlling for inhibition actually increased the strength of association between age and the TLC-E (β = - 0.76, p<0.001). This effect is due to the inhibition measure having a strong relationship with age, but a much weaker relationship with the TLC-E. It is likely that with increasing age one's ability to inhibit irrelevant information diminishes, but decreases in the efficiency of inhibitory processes bare little influence on the persons' performance on the TLC-E. This is not to say that inhibition is not a potentially important mediator of cognitive performance in older age, but that the skills assessed by the TLC-E are unlikely to require the involvement of inhibitory processing. The model indicates that inhibition is a suppressing factor, as it has a small correlation with the criterion variable but is strongly correlated with the predictor variable. Based on these considerations inhibition was not included in any further mediator analyses.

Figure 8. Path diagram illustrating relations among age, TLC-E and long-term memory.

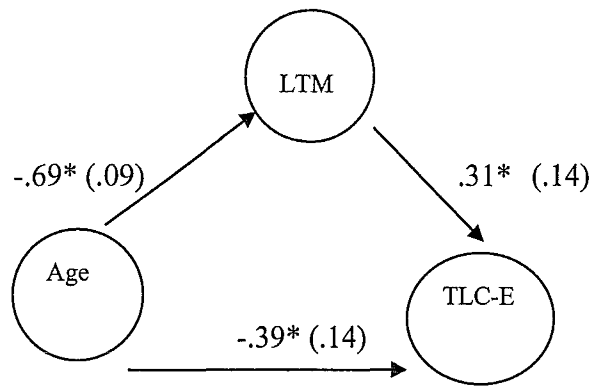


Figure 8 represents the path model including age, LTM and TLC-E. LTM demonstrated a significant moderate association with the TLC-E ($r = 0.58$, $p<0.001$) which remained significant after controlling for the effects of age ($\beta = 0.31$, $p<0.05$). Control for the effects of LTM also accounted for a proportion of age-related variance on the TLC-E, but the direct age-TLC-E path coefficient, that is after controlling for the effects of LTM, remained ($\beta = - 0.39$, $p<0.01$). In comparison to the other models LTM was comparable to speed in accounting for age-related differences on the TLC-E, but less successful than working memory.

In summary, the simple mediational models suggested that the association between age and TLC-E composite was mediated through an age-related decline on working memory. Speed and LTM also mediated the age-related effects to a marked degree, but only partially, as they both failed to remove the direct association between age and TLC-E. Inhibition was unlikely to contribute to age-related differences in language performance.

The next question of interest is whether working memory has an independent influence on cognitive functioning even when one controls for other potential mediators. As explained in the introduction, one popular suggestion is that age related differences in cognitive functioning are mediated only partly by working memory, which in turn is mediated age-related reductions in processing speed, and that speed itself has direct (non-mediated) effects on cognition (Salthouse, 1991). According to Salthouse (1980, 1991) the same pattern of relations should be detected for a measure like the TLC-E composite as he found for measures of cognition, since he asserts that speed "must be the principal mechanism behind age differences in nearly all aspects of cognitive function" (1980, p.61). Salthouse's model is represented in Figure 9a and the test of this model based on the current study is represented in Figure 9b.

Figure 9a. A hypothesized path model of relation among age, speed, working memory and cognition according to Salthouse (1991).

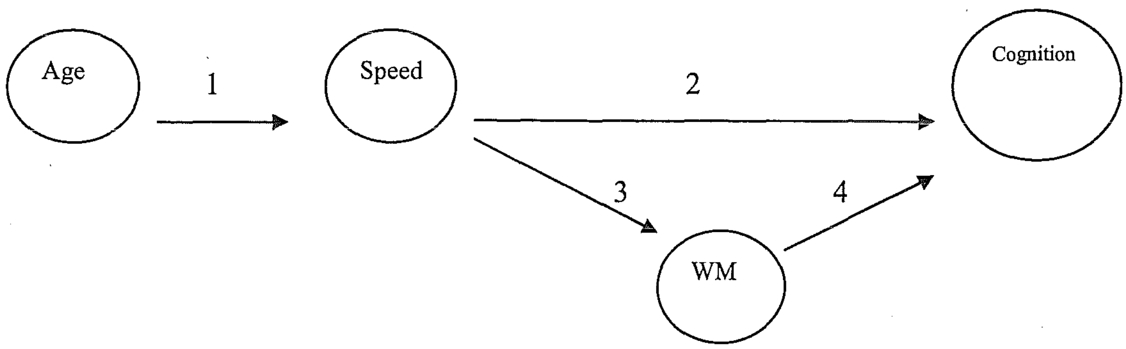
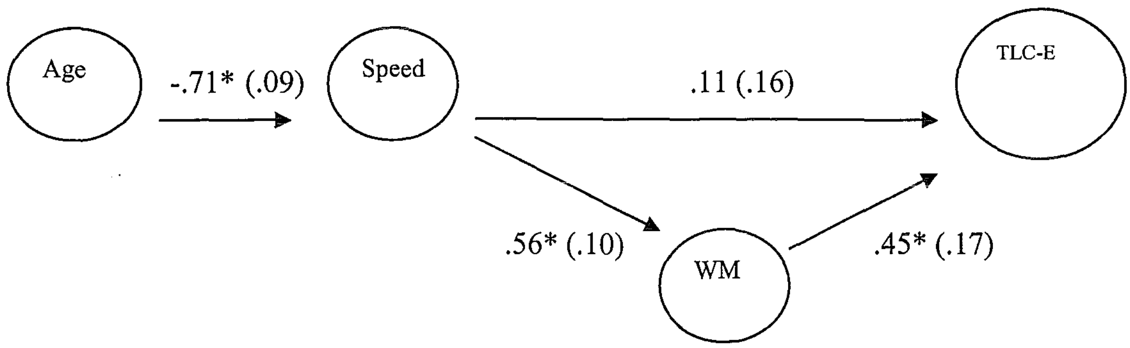


Figure 9b. Path diagram of relationship between age, speed, working memory and TLC-E in the present study, when Salthouse's (1991) model is tested.



The direct association between age and TLC-E (r= -.61; not shown) in Figure 9b is effectively abolished (β = -0.19 n.s.), but this model has two problems. As predicted by Salthouse (1991), speed mediated the relationship between age and working memory (path 3: β= 0.56, p<0.01). In turn, however, working memory mediated the association between speed and TLC-E (path 4: β= 0.45, p<0.01). Importantly, the independent effect of speed on TLC-E did not remain (path 2: β= 0.11, n.s.) after controlling for the effects of working memory and age. In addition, when one controls for working memory, the path coefficient between age and speed was β= -0.24, p<0.05 (not shown on Figure 9b), which is less than the path coefficient between age and working memory when one controls for speed (β= -0.34, p<0.01; not shown on Figure 9b). According to Salthouse (1991), this latter pattern should be the reverse.

The partial failure of Salthouse's model to explain age-related variance in TLC-E performance indicates that an alternative model, based on that predicted by Van der Linden et al (1999), warrants evaluation (Figure 10a). In this model, the direct path between speed and TLC-E is removed, but the path between age and working memory is included. The path coefficients obtained when this model is tested are presented in Figure 10 b. Again, the direct association between age and TLC-E in Figure 10b (r=-.61, not shown) is effectively abolished (β = -0.19, n.s.). As before, speed does not affect the relationship between age and TLC-E after controlling for the effects of working memory.

In this case, age and speed each independently influence working memory, which in turn uniquely mediates the relationship between age and TLC-E ($\beta = 0.45$, $p<0.01$). Speed partially mediates the relationship between age and working memory. That is, speed only indirectly affects the relationship between age and TLC-E through its effects on working memory. Working memory fully mediates relationship between age and TLC-E, beyond any indirect mediation by speed, because of the significant direct (i.e. independent) association between age and working memory.

Figure 10a. Path diagram of hypothesized relationship between age, speed, working memory and the TLC-E, based on Van der Linden, et al. (1999) model.

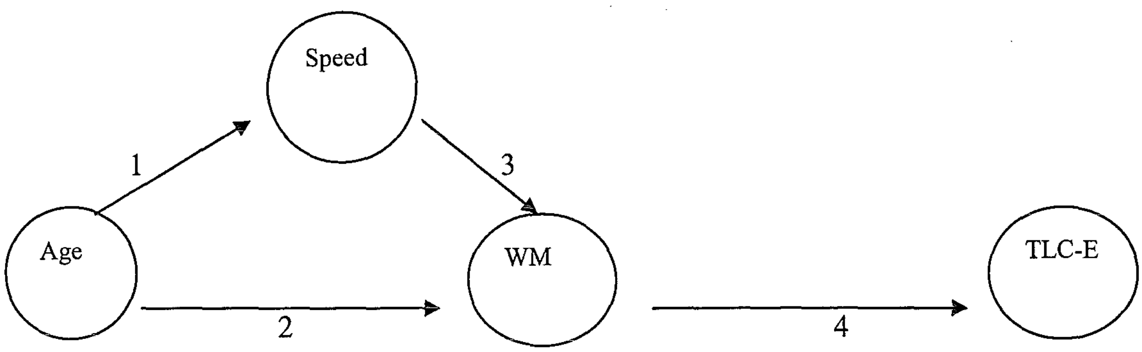
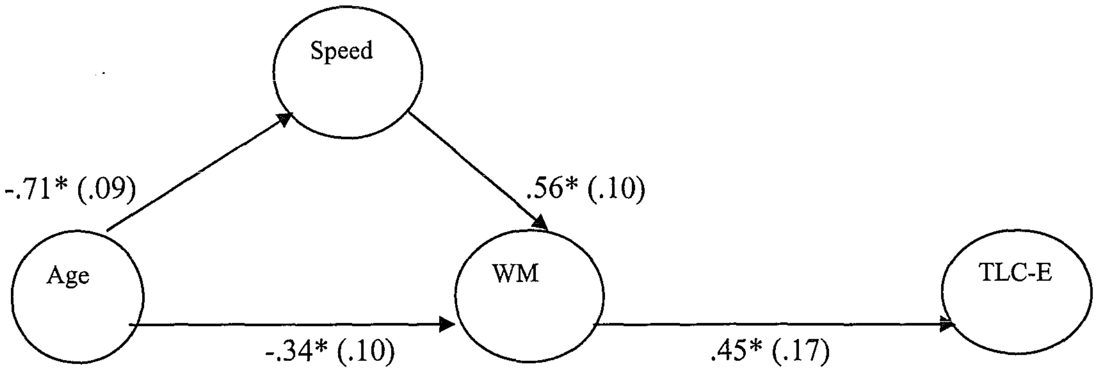


Figure 10b. Path diagram of relationship between speed, working memory, age and the TLC-E, when Van der Linden et al. (1999) model is tested



Previously, it has also been suggested that the TLC-E performance might be affected by an age-related decline in long-term memory. To control for the possible contribution of long-term memory, the model presented in Figure 10b was reassessed after any associations with LTM had been partialled out. The model provided in Figure 10b was maintained, because the obtained pattern of results did not change after controlling for LTM. Specifically, age was still independently related to working memory (β=-0.24 (.10), $p<0.05$) (note, that standard errors are provided in brackets) and the independent relation between working memory and TLC-E was maintained (β=.43 (.18), $p<0.05$). The direct relationship between age and TLC-E (β=-.18, (.147), n.s.) remained negligible. The association between speed and working memory was maintained (β=.42 (.11), $p<0.01$). The LTM measure demonstrated a moderate degree of association with working memory (β=.29 (.10), $p<0.05$). The independent contributions of LTM to TLC-E (β=.05, (.15), n.s.) and speed to TLC-E (β=.10 (.17), n.s) were negligible. Thus, even after controlling for the effects of LTM, working memory still uniquely mediated the effects of age on the TLC-E. Age, speed and LTM had an indirect effect on the TLC-E through their effects on working memory.

Path models for the TLC-E subtests

As the TLC-E composite consists of subtests that measure diverse language functions the nature of the relationship between the individual subtest that the mediator variables was also examined. Zero order correlations between the TLC-E subtests and the composite measures used in path analysis are presented in Table 37. A path model depicted in Figure 10a was tested, however instead of the TLC-E composite, each individual TLC-E subtest featured as a dependent variable. The obtained pattern of results was then re-examined after partialling out the effects of LTM. The path coefficients for model in Figure 10a (for each subtest) are presented in Table 38. Table 39 provides path coefficients after controlling for the effects of LTM. It should be noted that the magnitude of the direct relationship between age and LTM (β=-.69 (.09), $p<0.001$) and between LTM and working memory (β=.29 (.10), $p<0.05$) remained the same for each subtest (as it was the only variable manipulated) and hence is not cited in the table. For completeness Tables 38 and 39 also provide coefficients for the direct links between age and the TLC-E subtest and speed and the TLC-E subtest, as it was hypothesized that the magnitude of these direct associations would be negligible, the paths were not included in the diagram in Figure 10a.

Ambiguous Sentences

As evident from the Table 38, the path model depicted in Figure 10a provided a good representation of associations between the mediator variables and the Ambiguous Sentences subtest. That is, age and speed both affected Ambiguous Sentences performance through their effects on working memory. The working memory was uniquely and strongly associated with the Ambiguous Sentences performance. The independent associations between speed and Ambiguous Sentences and age and Ambiguous sentences were negligible.

Table 37. Zero order correlations between TLC-E individual subtests and mediator varibles in the path analysis.

	Working Memory	Speed	LTM	Age
Ambiguities	.70	.66	.61	-.58
Inferences	.64	.63	.52	-.61
Recreating sent.	.78	.64	.58	-.72
Figurative	.45	.40	.48	-.24♦

Note: ♦p >0.05, all other *ps* <0.02

Table 38. Path coefficients and (standard errors) for the relationship between age, working memory, speed, and individual TLC-E subtests.

	PATH					
	1	2	3	4	5	6
Ambiguous Sentences	.-71	-.34	.56	.43	.25♦	-.09♦
	(.09)	(.10)	(.10)	(.17)	(.16)	(.14)
Making Inferences	-.71	-.34	.56	.27♦	-.23♦	-.23♦
	(.09)	(.10)	(.10)	(.18)	(.17)	(.15)
Recreating Sentences	-.71	-.34	.56	.59	-.08♦	-.34°
	(.09)	(.10)	(.10)	(.14)	(.13)	(.12)
Figurative Language	-.71	-.34	.56	.49°	.17♦	.24♦
	(.09)	(.10)	(.10)	(.21)	(.21)	(.18)

Note ♦p>0.05, °p <0.05, all other *ps* <0.02. Path 5 and 6 are not depicted in the diagram in Figure 10a, path 5 is for the direct relationship between speed and the TLC-E subtest, path 6 is for the direct relationship between age and the TLC-E subtest.

Table 39. Path coefficients and (standard errors) for the relationship between age, working memory, speed and individual TLC-E subtests after controlling for the contribution of long-term memory.

	PATH					
	1	2	3	4	5	6
Ambiguous Sentences	-.38	-.24	.42	.39	.22♦	-.07♦
	(.11)	(.10)	(.11)	(.18)	(.16)	(.14)
Making Inferences	-.38	-.24	.42	.29♦	.25♦	-.25♦
	(.11)	(.10)	(.11)	(.19)	(.17)	(.15)
Recreating Sentences	-.38	-.24	.42	.65	-.04♦	-.37°
	(.11)	(.10)	(.11)	(.15)	(.14)	(.12)
Figurative Language	-.38	-.24	.42	.32♦	.07♦	.33♦
	(.11)	(.10)	(.11)	(.22)	(.20)	(.17)

Note ♦p>0.05, °p <0.05, all other *ps* <0.02. Path 5 and 6 are not depicted in the diagram in Figure 10a, path 5 is for the direct relationship between speed and the TLC-E subtest, path 6 is for the direct relationship between age and the TLC-E subtest.

The same pattern of results was maintained after partialling out the contribution of long-term memory, with working memory continuing to demonstrate strong associations with the Ambiguous Sentences. The direct contribution of the LTM to the Ambiguous Sentences was negligible (β=.10 (.15), n.s.)

Making Inferences

The proposed path model (Figure 10a) could not adequately explain the individual differences observed on the Making Inferences subtests. The path coefficients obtained for this subtest indicated that neither working memory, speed, nor age were successful in accounting for the observed differences on this subtest. Partialling out the long-term memory did not have any additional effects on the observed associations. The long-term memory also did not exhibit an independent association with the subtest (β= -.06 (.16), n.s.).

Recreating Sentences

The path model depicted in Figure 10a provided a good representation of the associations between the mediator variables and the Recreating Sentences subtest. However, there was evidence of a direct, unmediated link between age and the subtest. The effects of speed and long-term memory on this subtest were mediated through the effects on working memory, with working memory demonstrating a strong independent association with the subtest. The magnitude of the direct association between LTM and the subtest was negligible (β=-.15 (.12), n.s.). Age had both direct and indirect (through working memory) effects on the Recreating Sentences subtests that were maintained after controlling for the LTM contribution.

Figurative Language

The path model in Figure 10a served as a good representation of the relations between the Figurative Language and the mediator variables when the effects of working memory, speed and age were examined. However, when the effects of LTM were partialled out the model was no longer adequate. Accounting for the effect of LTM resulted in reduction of the working memory-Figurative Language association to non-significant levels.

However, the long-term memory demonstrated a strong independent association with the subtest (β= .40 (.15), $p<0.05$), suggesting that individual differences observed on the Figurative Language are best accounted for by the differences in long-term retention.

In summary the path analyses for each individual TLC-E subtest indicate that processing speed did not have independent associations with any of the subtests. The relationship between speed and individual subtests was fully mediated by working memory. Working memory had a unique contribution to the performance on the Ambiguous Sentences and Recreating Sentences. The individual differences in Figurative Language abilities were best accounted for by the differences in long-term memory efficiency. None of the proposed mediating factors were successful in accounting for the presence of individual differences on the Making Inferences subtest.

4. DISCUSSION

One goal of the present analysis was to assess the performance of older adults on the Test of Language Competence-Expanded Edition (Wiig & Secord, 1989) and to develop provisional normative data for the New Zealand elderly on this test. Hence, this section first addresses the issue of the representativeness of the present sample, as it is pertinent to the interpretation of the results obtained and particularly to norm development. Next, the general age-related findings produced by the TLC-E and other cognitive measures are discussed. The second goal of the study was to evaluate the effects of age-dependent variables on age differences in discourse abilities, so the factors that appear to mediate language performance are then addressed. Specific emphasis is placed on reviewing the nature of the independent contribution of working memory to the TLC-E performance. The TLC-E is comprised of a number of subtests that measure various language functions, so issues relevant to each individual subtest are also reviewed. Finally, the major contributions and limitations of the current study are summarized.

4.1 Representativeness of the Sample

The present study succeeded in testing two representative samples of people over the age of 65, young-old (65-74 years) and old-old (75-89 years), together with a comparative reference group aged 20-34 years. These volunteers were recruited to ensure they constituted a cross-section of general population. There was a good spread of ages in each group and close to equal number of males and females. A specific emphasis was placed on recruiting participants of differing levels of education. The samples were stratified according to the age-appropriate educational level distribution of the New Zealand population (Statistic New Zealand, 1996). As reflected in Table 4 a close correspondence between the samples and the New Zealand population proportions according to age, sex and education was obtained. This strategy ensured cross-sectional comparability and the representativeness of the current samples relative to the local community. Attempts to recruit adults of differing educational levels also ensured that the participants came from a broad community base rather than just from the University population or University volunteer database.

All participants were generally healthy, and there was no sign of failing cognitive status (as measured by the MMSE) or depression (as assessed by the BDI). The elderly groups in particular were comprised of individuals living independently and not suffering from any non-corrected vision or hearing impairment to such a degree that their performance on cognitive tasks could have be compromised. At the same time the study was careful to avoid recruiting a sample of "super healthy" elderly, following the suggestion that norms for the elderly people need to be representative of the health profile of the elderly population and include participants with mild medical illnesses (e.g. medication-controlled hypertension) but exclude those with active neurological or psychiatric disorders (Malec, Ivnik, & Smith, 1993). That is, the older individuals who had well-controlled hypertension, arthritis or rheumatism were included in the present sample.

Naturally, the present sample had some limitations that could have potentially affected its representativeness. Firstly, the sample size per age group was relatively small; the danger of the small size is that performance of individuals in that sample may not adequately represent the range of abilities in the general population. While it would have been desirable to obtain a larger sample size, the time involved per participant (nearly 5 hours for each person) was considerable and the total sample size an adequate compromise. Nonetheless, the wide distribution of educational qualifications in the current sample counteracted some of the negative effects of the small sample size by ensuring adequate variability in scores, which was clearly achieved. Another feature of the sample that can be seen as a possible threat to its representativeness is the mildly higher than average intellectual ability of the participants. The tendency of the New Zealand participants to obtain higher than average IQ scores relative to the American population has been demonstrated in the past (see Harvey & Siegert, 1999). That is, while higer IQ scores may represent a form of sampling bias in the current study, it is equally plausible that the American norms available for the WASI (and possibly the WAIS-III) underestimate the abilities of the New Zealand population.

Another issues it that the elderly in the sample were all living independently and were perhaps more physically active participants in their community and somewhat healthier than an average sample. The study did not obtain any information with regard to the amount of community support the older adults were receiving (i.e. whether they were semi-dependent). The level of dependency has previously been shown to be associated with the effectiveness of cognitive functioning (Harvey & Siegert, 1999). Also, while an attempt was made to sample participants from a broad geographic catchment area within the greater Christchurch region and from suburbs thought to represent different socio-economic sectors of the community, we were mostly reliant upon finding volunteers. Thus we must also acknowledge possible problems with the generalizability of the findings to older adults from other New Zealand regions. Lastly, the present study did not directly obtain the information on the cultural/ethnic background of the participants. To the question of: "What language do your speak at home?" only two of the participants stated English and Maori, suggesting that they were of Maori descent. No other information with regard to the ethnic composition of the sample is available.

In summary, the sample obtained was comprised of 42 New Zealand elderly living independently in the community and 20 young adults. Although the representativeness of the sample has some limitations, it adequately reflected the range of educational backgrounds of the constituent age groups. A good representation of New Zealand population educational proportions was a particularly important feature of the sample, which also ensured appropriate variability in test scores obtained. In short, the current sample may be accepted as a standard against which an adult individual who is being assessed may be measured.

4.2 Performance of Elderly on the Test of Language Competence-Expanded Edition

One important new finding that this study has provided is that older adults exhibit clear age-related deterioration on the Test of Language Competence-Expanded Edition, a test of higher language functioning. Both elderly groups performed substantially (as much as one standard deviation) below the mean level of the younger adults on the TLC-E discourse composite. The differences in performance could not simply be attributed to differences in general cognitive functioning (as measured by the MMSE), intelligence or verbal knowledge, since tasks that measure these functions yielded no group differences and covariation for the measures did not remove the observed age differences. While the young and the old-old groups differed in years of education, the group effects on the TLC-E were also maintained after controlling for the effects of the years of education. There were no gender differences on the TLC-E. As the TLC-E composite is thought to reflect general higher language competence and discourse skills, it can be concluded that these abilities are impaired in the elderly in comparison to the young. Moreover the old-old elderly scores dropped significantly in comparison to those of the young-old. This finding provides evidence of increased vulnerability of the old-old adults to impairment in discourse skills and argues against the stability of higher language function in the elderly population

4.3 Clinical Utility of the TLC-E

As the TLC-E demonstrated particular sensitivity to age-related changes in higher language function it warranted the development of age specific norms. The present research has provided provisional norms for the TLC-E for the reference group and the two elderly groups in the age bands 65-74 and 75-89 years. The normative data provided presents a valuable point of comparison for future research on changes in language skills associated with different neurological conditions. The present data clearly indicated that the TLC-E is a sensitive enough instrument to detect deterioration not only in the elderly in comparison to the young, but also within the elderly group. This sensitivity of the TLC-E to decline, a characteristic not commonly shared by the standardized aphasia or neuropychological battery tests, has already been recognized as advantageous in assessment of language deficits associated with Multiple Sclerosis (Lethlean & Murdoch, 1997), Parkinson's disease (Lewis et al, 1998) and Dementia of the Alzheimer Type (Harris, 1994). Given the TLC-E sensitivity to normal age-related decline and previous reports of the TLC-E's ability to differentiate between cases of moderate and severe dementia (Harris, 1994), the test can be a useful tool in identifying early, pre-clinical cases of dementia. It is a well-documented finding (Kemper & Kemtes, 2000) that in its severe form dementia of Alzheimer's type is characterized by impairment in discourse language, with discourse of the patients being tangential and incoherent. However, in cases of mild dementia discourse difficulties are not as apparent, and much more difficult to observe. The normative data provided by the present study, particularly that for the old-old group performance, may serve as a valuable benchmark against which the performance of the patients can be compared, and may provide an extra source of diagnostic information.

The normative information collected in the present study may also help to provide an accurate and meaningful neuropsychological assessment of older New Zealanders. With New Zealand's older adult population growing, and the psychometric testing of older adults becoming more common, there is an increased demand for tools that are not only sensitive enough to detect any decline in cognitive abilities, but also have age-appropriate norms. The fact that the currently provided norms for the TLC-E have an age range that goes up to 89 years increases the utility of the instrument for use with the elderly population. Importantly, we have found the participants to respond positively to the TLC-E. The administration of the test actually aided substantially in building rapport, unlike other standardized tests (e.g. the WMS-III subtests). Clinicians would also find the TLC-E easy to administer and user friendly.

An additional value of the TLC-E is that it assesses language in a context that closely approximates natural discourse. Hence, the results of the assessment can be used to provide valuable information to the support people and caregivers on how to modulate the use and form of their speech based on the actual communicative needs of the older person.

Evaluation of the TLC-E reliability suggested that the composite score and three out four subtests had robust internal consistency. Good internal consistency of the TLC-E indicates that the presence of American expressions in the test, as well as the holistic scoring rules (rules that rely, in part, on the examiner's judgment of accuracy and acceptability of the utterance), did not substantially affect the reliability of the results obtained. The comparison of the current reference group performance to that of the 18+ year olds detailed in the TLC-E manual produced relatively small differences only, supporting the notion that comparable results can be obtained on this test with the New Zealand young adult population. Some caution is warranted, however in interpreting the scores on the Figurative Language subtest. The present study found this subtest to be less sensitive to age-related decline and to have poor reliability.

In summary, the information collected in the present study is useful for future evaluations of the TLC-E in related studies, experimental research and clinical practice. Moreover the obtained results provide valuable insights into the nature of age-related decline in discourse skills. The normative data obtained in the present study presents a major step towards the development of an adequate local normative database for older people.

4.4 Older Adults' Performance on Working Memory Measures

The present study examined the performance of younger and older adults on a number of different working memory measures in attempt to better capture the underlying construct. As expected, the results of this study augment an extensive literature demonstrating that older adults perform less well on tasks presumed to measure the storage and processing capacity of working memory (Babcock & Salthouse, 1990). Older adults exhibited clear decrements in their ability to divide resources between sentence comprehension and remembering final words, as reflected in the Daneman and Carpenter reading span performance. Interestingly, the decline in the reading span scores were most notable from young to young-old with little decrement evident with further age, which is consistent with previously reported findings (Merugo et al., 2001, Waters & Caplan, 2001).

The WMS-III working memory subtests also exhibited sensitivity to age-related decline. Deterioration in performance on the Letter/number Sequencing subtest, which assesses the ability to order and organize the information in working memory, was age progressive, with the old-old elderly exhibiting more decline than the young-old. Spatial Span performance, which relies on maintaining the memory for spatial pointing movements and reflects the operation of the visuospatial sketchpad, was also compromised in the elderly, although there was no evidence of the old-old being further impaired than the young-old. Digit Span has traditionally been used as a measure of working memory (see Light, 1990), but in the current study the total score demonstrated poor sensitivity to age-related deterioration. Age decrements were noted on the Digits Forwards subtest, but this task mainly assesses the operation of short-term memory retention and fails to address the working memory capacity for actively processing the information. On the contrary, Digits Backwards, which is generally regarded as a measure of working memory ability to re-organize information, failed to detect any age effects. Previous research (Babcock & Salthouse, 1990, Verhaeghen, et al., 1993) has questioned the validity and reliability of the Digit Span as a measure of working memory, and current results certainly indicate that it is insufficient to just rely on this task when an estimation of working memory functioning is required. Digit Span seems to provide an extremely narrow (if any) assessment of working memory construct, and fails to adequately capture the information-manipulation capabilities of the system.

The last working memory task used in the present study was semantic fluency. The older adults' capacity to generate items from a semantic category evidenced substantial decline in comparison to that of the young. Consistent with Engle's (1996) observation the young adults performed only slightly better than the elderly groups in the first minute of the five minutes of the task. As the time progressed, the differences in number of words generated started to increase, when the younger adults continued to perform better than the elderly and the young-old outperforming the old-old. Notably, this observed differential increase in the number of words generated could not be simply attributed to age differences in the rate of change in word production over time, as the groups did not differ on this measure. It is more likely that the main determinant of age differences on the fluency task was the number of between cluster switches made by the participants. Switching represents an ability to get access to a label of a new cluster and requires one to engage an in effortful, controlled search of the category. Once the cluster label has been accessed the items in the cluster become activated relatively automatically. The younger adults in the present study demonstrated a superior ability to switch between categories in comparison to the elderly, and the young-old were in turn superior to the old-old. On the contrary, the number of items recalled per cluster did not differ between the young and the old. Thus, the time between clusters was shorter for the younger adults (as reflected in higher number of switches), but once the recall of a cluster had begun, the inter-item time was nearly the same for both younger and older adults (as reflected in cluster size). According to Engle (1996) the number of words produced and switching on the fluency task reflect the controlled sustained attention properties of the working memory system, so one can infer that the older adults have insufficient attentional resources that compromised their performance.

The results of the present study suggest that the age-related influences on different working memory measures overlap to a large degree, which is supported by the finding of good interrelationships between most of the working memory measures examined here. All working memory measures, with the exception of clustering and rate of change on the fluency task, demonstrated a substantial amount of commonality as reflected in the robust reliability coefficient of the working memory composite.

4.5 Older Adults' Performance on the Processing Speed Measures

As expected the present results also support the notion of cognitive slowing with age. The older adults showed a pattern of progressive decline on tasks of information processing (Digit Symbol) and color naming. The fact that older adults tend to slow down on the Digit Symbol and color naming has been consistently demonstrated in the past (Kwong See & Ryan 1995, Salthouse, 1996, Van der Linden, et al., 1999). Word reading speed evidenced a decline only in the old-old adults. The observed differential sensitivity of the speed measures to age-related decline is actually consistent with the processing speed theory. Salthouse (1996) asserted that age-related affects could be expected to vary in magnitude on different speed measures, because of the operation of other influences, even when a common underlying mechanism is involved. For example, the degree of how well rehearsed (automated) a certain task is will effect the speed of performance. Slowing with age is often considered to be one of the best-documented and least controversial behavioral phenomenon in aging (see Salthouse, 1996). The median correlation between age and measures of speed is normally r= -0.45 across the life span (Salthouse, 1996). The correlations for age-speed association ranged from -0.43 to -0.77 in the present study. It is, however, unclear what mechanism might underlie slowing in old age. Salthouse (1980) suggested that the central nervous system may be functioning at a slower rate in older adults, but physiological studies have failed to find slowing with aging in latency of the cortical evoked potentials (Bashore, Osman, & Heffley, 1989).

4.6 Older Adults' Performance on the Inhibitory Efficiency Measures

Several aging studies, using both the traditional Stroop task and paradigms designed to produce Stroop-like measures (negative priming), have documented that older adults are more susceptible to interference (Kwong See & Ryan, 1995, Rogers & Fisk, 1991, Van der Linden et al, 1999). The interpretation of this finding has been to associate older age with poorer ability to inhibit irrelevant information that competes for the control of behavior. Consistent with previous findings progressive age-related declines were also observed here in the ability to name incongruent color words. The present study also computed a Stroop interference score which controls for the speed of reading and color naming. The older adults showed more interference than the young adults on this measure, but there was no evidence of deterioration with further age, (with the old-old adults performing nearly as well as the young-old). As the majority of previous of aging studies (e.g. Kwong See & Ryan, 1995, Van der Linden, et al., 1999) do not adequately control for the speed factor, the greater Stroop interference effects found in the elderly samples may not represent an accurate reflection of age differences in inhibitory efficiency.

4.7 Older Adult's Performance on the Long-term Memory Measures

Performance of the elderly on the WMS-III delayed memory tasks confirmed the expectation that performance on these tests declines with age. Progressive deterioration was noted for the ability to remember and later recall a pair of semantically unrelated words (Verbal Pairs task). This result is consistent with the previous finding of paired associate learning being highly sensitive to age decline (Kausler 1994). The performance of elderly on the Logical Memory subtest also indicated the presence of age-related deterioration. This decline was not progressive however, as the old-old adults performed at virtually the same level as the young-old. The differences in age sensitivity of the tasks can be attributed to different processing demands imposed by the tasks. Logical Memory provides the individual with contextual details, making the information easier to encode and retrieve. On the contrary, the verbal pairs require the individual to organize the material more actively to enable encoding and subsequent retrieval, hence leading to more profound impairment.

4.8 Performance of the Current New Zealand Sample on the WMS-III

Administration of a number of the WMS-III subtests provided the current study with an opportunity to review how New Zealand elderly and the reference group fare on this test in comparison to the American normative population. The old-old group in the present sample performed at comparable levels to the American normative sample on the Working Memory and Auditory Memory Indices. These data can be taken as preliminary indication of the reliability of the subtests that comprise these indices for the 75-89-year-old New Zealand elders. Contrary to this finding, significant differences were detected between the WMS-III population mean and the current sample mean for age groups 20-34 years and 65-74 years which applied to the Working Memory and Auditory Memory Indices. It should be noted that the differences were relatively mild and never exceeded one standard deviation. Nevertheless these differences are of clinical significance and may potentially result in underestimation of memory decline in individuals suffering from neurological conditions. Previous reports of standardization of other neuropsychological instruments (e.g. the Rivermead Behavioral Memory Test (Frazer, Glass, Leathem, 1999), the Graded Naming Test and the Recognition Memory Test (Harvey & Siegert, 1999) on a New Zealand elderly population suggested that the elderly tend to perform slightly better than the American or British cohorts. One of the explanations provided for these differences was varying educational levels of the New Zealand and American samples (Harvey & Siegert, 1999). It is, however, less likely that years of education contributed to the presently obtained differences, as the mean number of years of education of the present sample (young-old - 11.4 years and young - 12.8

years) nearly equaled that of the American population (young-old - 11.7 years and young - 13 years). Slightly higher than average IQ level of the participants in the present sample could have potentially been responsible for the observed differences, as higher IQ has been cited in the past as one of the protective factors against memory decline, especially for the younger elderly (Hultsch, et al., 1999). It should be noted that the limitations of the current sample might have also contributed to the discrepancies obtained between New Zealand sample and American population. However, the findings in the present study do provide an important initial step, however, towards developing New Zealand norms for the WMS-III, and indicate an urgent need for more comprehensive normative investigations on this widely used instrument.

4.9 Cognitive Mediation of Age Differences in Discourse Functioning

The previous section detailed that many moderate changes occur in the processing skills in the elderly. One of the aims of the current study was to establish how age differences on discourse composite are mediated by age-related declines in these information processing skills. The main hypothesis was that if any of the postulated factors indeed mediate age-language relationship then any direct age-related effects would be expected to be reduced, when the variation associated with the hypothesized mediator was eliminated.

The results of the current study can be summarized as follows. First, all significant relationships between age and the discourse composite were indirect and mediated through age-related reductions in speed, long-term memory and working memory. These findings indicate that the three general factors of cognitive function may be useful in explaining age-related differences in higher language skills. Furthermore, the results indicated that the contribution of speed and long-term memory to language performance was indirect and was mediated through working memory. Direct mediation of the association between age and language skills was also found.

With regard to the relationship between language functioning and working memory, the present results are consistent with the view that working memory plays a critical role in discourse processing (Stine, 1995). Stine's core assumption was that the function of working memory in discourse processing is to recode the verbatim representation into a semantically-based form. A consequence of this view is that the older adults' limited working memory capacity affects their ability to process discourse, by limiting the number of propositions that can be processed at a time. Age decrements should be particularly notable when the disparity between the linguistic input and contents of the knowledge base are substantial and the older adult has to make extensive transformations of the input to make some connection with the knowledge (e.g. when interpreting ambiguous words).

The findings are only partly consistent with the view that speed of processing variable exerts part of its effects on language processing through working memory (Salthouse, 1992). In addition to the indirect contribution of age to working memory (via speed), there was a direct and strong link between age and working memory. Thus, the age-related differences in discourse function were explained by a reduction of the capacity of working memory which was itself partly influenced by the reduction of speed. This particular relationship between age, working memory and discourse processing is similar to that observed in Van der Linden et al. (1999) study. It is not totally contradictory to the Salthouse's (1996) processing speed theory, but Salthouse draws very strong conclusions that the primary mediator is cognitive slowing which also affects working memory, and hence impairs cognitive performance (Salthouse, 1996). Yet he has also pointed out that the effects of slowing are not necessarily universal for all cognitive measures. Although seldom emphasized, Salthouse (1990) found that a working memory contribution to cognitive performance was over an above that of speed when cognitive functioning was assessed using accuracy of responding rather than speed of responding. Previous studies that have found discourse processing-speed associations have also normally used language tasks that relied on fast responding (Stine & Wingfield, 1987, Stine, et al., 1986, Tun et al., 1982). As the TLC-E de-emphasizes the speed of responding, the obtained pattern of results can be seen as a more accurate reflection of cognitive skills' contribution to discourse functioning.

The present results also indicated that speed exerted its influence on the TLC-E through working memory, which is in concert with the view that speed partially affects working memory function. After reviewing a number of lines of evidence Salthouse (1994) concluded that speed's influence on working memory occurs because older adults are slower than young at encoding information or establishing adequate internal representation, but not because of age differences in the rate of which the information is lost. Given this conclusion, it would be fruitful for the future research to examine the extent to which speed is involved in various working memory tasks. There is also a need for an agreement on what constitutes proper measurement of speed and working memory. It seems clear from the results of the present study that strong associations obtained between working memory and discourse function were in part due to the range of working memory measures selected. There is a convincing evidence that the central executive (the supervisory component of working memory) is likely to be a cluster of several control processes (Baddeley, 1996). Thus, measures of working memory that tax several of these processes should be better at predicting cognitive performance. It is likely that previous failures (Kwong See & Ryan, 1995) to detect independent working memory-language associations were due to the circumscribed choice of working memory measures. For example, Kwong See & Ryan only used digit span and digit lag as measures of working memory, and hence, have likely failed to capture all the aspects of the underlying construct.

The finding of strong independent contribution of working memory to discourse processing was further confirmed when the effects of long-term memory were partialled out. A direct link between age and working memory remained, but age also affected working memory indirectly through long-term memory. This finding is consistent with the general view of LTM-working memory relationship. Ericsson and Delaney (1999) have previously asserted that ease of encoding and retrieval of information in LTM contributes to the efficient performance on working memory tasks, which in turn may modulate the observed age differences on a criterion measure. As older adults demonstrate more impairment in the ability encode/retrieve information they should demonstrate more impairment on working memory tasks, which are more reliant on these functions. Indeed, in the present study reading span (performance on which is greatly assisted if the individual is able to form associations between words, i.e. actively encode) and the total number of words recalled on the fluency task (that reflects retrieval efficiency) demonstrated strong associations with the LTM composite.

Long-term memory itself however, did not demonstrate direct independent associations with the discourse composite. Its contribution to the composite was indirect, through working memory. This is somewhat contradictory to what has been reported by Singer and Richot (1996) who found the LTM and working memory to contribute independently to discourse processing. These differences may be because Singer and Richot did not employ statistical control procedures like path analysis and confided their analysis to obtaining simple correlations. Additionally, they used a different measure of long-term memory, an integration task, which more heavily draws on reasoning and retrieval abilities rather than retention.

The current findings challenge the assumption that age differences in the inhibition of irrelevant material underlie age differences in discourse processing (Hasher & Zaksc, 1988). The present study did not find evidence for inhibition mediated age-related differences in language functioning. Inhibitory efficiency was only very weakly related to the performance on the TLC-E composite, suggesting that the discourse functions captured by this task do not heavily rely on the ability to inhibit irrelevant information. The current finding is in concert with what has been previously reported. For example, Salthouse and Meinz (1995) after examining the association between age, working memory and inhibition, found that removing the influence of inhibition on age did little to reduce the association between age and working memory. Grant and Dagenbach (2000) not only failed to find age differences in negative priming (a measure of inhibition), but also any evidence of association between inhibition and discourse processing, suggesting that the influence of inhibition is independent of age differences on language tasks.

By contrast, two studies by Van der Linden et al. (1999) and Kwong See & Ryan (1995) found Stroop interference to mediate age differences on language tasks. A possible explanation of this discrepancy is that different approaches to assessment of inhibition resulted in different findings obtained with regard to the strength of discourse-inhibition association. There is no agreement between researchers on how inhibition and interference should be measured. Even when the Stroop task is employed, different variants of the Stroop are often used (for e.g. Salthouse and Meinz used reading words, reading colors, naming quantities and naming positions as their measures of interference). Additionally the observed Stroop effects are often not corrected for speed (e.g. Kwong See & Ryan (1995) and Van der Linden et al. (1999) research) leaving open the possibility that it is the speed factor that potentiates the observed language-interference relationship. From the theoretical standpoint, it also is unclear how interference score is actually related to the concept of inhibition. In the interference condition inhibitory processes should intervene to reduce interference during concurrent response selection, but evidence for this interpretation is controversial (see Kieley & Hartley, 1997). Kramer et al., (1994) have found evidence of the relative independence of various inhibition measures, which suggests that if other measures of inhibitory processes (rather then Stroop interference) are used a hypothesized relationship between working memory and discourse skills may be supported. At present, the evidence for the inhibition mediated language performance remains controversial. An absence of clear relationship between inhibition and discourse processing found in present study suggests that either the original formulation by Hasher & Zacks (1988) may need modification or further elaboration, or that other measures of this construct might well be in order.

In summary, the present study showed that the construct of working memory, as assessed by a composite measure, was a good predictor of performance on the discourse task. Neither speed of processing nor long-term memory directly explained performance on this discourse measure. Resistance to interference only showed a very weak initial association with the higher language composite. These findings are compatible with the results of other studies. The present results obtained by cross-sectional analysis would require further confirmation using other research methodologies, like longitudinal design. It is worth noting that in their longitudinal study Hultsch et al. (1999) reported results very much inline with what was found here. In particular, in that study the change in working memory determined the changes in comprehension and memory, and the influence of speed was mediated through changes in working memory. Finally, it is important to point out that current results suggest that the extent to which different general factors contribute to cognitive performance depends on the type of cognitive task or the information to be processed.

4.10 The Contribution of Mediator Variables to Older Adults' Performance on Individual TLC-E Subtests

The TLC-E discourse composite is compromised of a variety of subtests that assess diverse language skills. Thus, the question as to how the component subtests may be affected by age-related changes is also of interest.

Performance of Elderly on the Ambiguous Sentences Subtest

The elderly demonstrated clear decline in performance on the Ambiguous Sentences subtest. The observed deterioration was also progressive, with the old-old adults performing significantly worse than the young-old. The Ambiguous Sentences subtest evaluates that ability to recognize and interpret the alternative meanings of selected lexical and structural ambiguities. It is thought that when lexical ambiguity is encountered all meanings are simultaneously activated upon presentation of an ambiguous word. Once the meanings of the word have been accessed the context influences the selection of the appropriate meaning (Onifer & Swinney, 1981). In the present study, the older adults were less able to generate or access the alternative meanings of an ambiguous word. For example, one older person's two similar interpretation of the lexically ambiguous sentence "*I knew that glare really bothered Jane*" were " *the glare from the sun* " or "*the glare from the light*". The interpretation of syntactic ambiguity has been a subject of debate in the literature, with limitations in processing resources often cited as an explanation for the observed age differences (Kemtes & Kemper, 1997).

The present study clearly demonstrated that age differences in processing resources mediated an age-related decline in Ambiguous Sentences. Working memory, speed and LTM were able to successfully account for most of the age-related variance on this subtest. As with the composite language measure, working memory directly mediated the effects of age, whereas the contribution of speed and LTM was indirect, through their effects on working memory. The finding that working memory had strong association with the subtest conflicts with Harris' (1994) finding where weaker associations were detected. One interpretation of this discrepancy is that the working tasks used in Harris's study (digit lag and digit ordering) might not have fully captured the working memory construct, thus resulting in the weaker associations obtained. Also, Harris's sample consisted of Alzheimer's dementia patients only, which could have been another contributing factor to the different pattern of associations obtained. Other studies that examined how the processing limitations affect ambiguity resolution have produced variable results. For example, MacDonald, et al. (1992) found individuals with low working memory capacity to experience significant difficulties in disambiguation. On the contrary, Kemtes & Kemper (1997) reported weak associations between on-line ambiguity resolution and working memory capacity. Working memory capacity, however, played a more significant role in modulating ambiguity comprehension when it was examined off-line. One of the problems with the off-line research findings is that the off-line tasks normally demand from the participant to retain large amounts of syntactically complex material in memory or interpret implausible sentences, thus placing unduly heavy demands on memory for the sentence presented. In contrast, in the present study, while the comprehension was assessed off-line, all of the ambiguous sentences were plausible, there was no demand for memorization of material and the ambiguity difficulty was only moderate. Thus, the present results favor an interpretation that working memory may underlie performance on ambiguity resolution tasks, even when the material to be processed is moderately difficult and the demand on memorization of information is minimal.

Performance of Elderly on the Making Inferences Subtest

Both of the elderly groups experienced significant difficulty selecting appropriate choices for inference resolution in comparison to the young adults on the Making Inferences subtest. The performance of the old-old group was even further impaired than that of the young-old. The Making Inferences subtest evaluates the ability to make permissible inferences on the basis of existing causal relationships or chains in short paragraphs. There is mixed evidence concerning the comprehension of inferences in old age (see Light, 1990). The debate revolves around the issue of whether older adults demonstrate genuine declines in inference comprehension (Till, 1985, Till & Walsh, 1980), or whether the observed declines are a product of the excessive memory load imposed by the task (Burke & Yee, 1984, Ligth, et al., 1980). The research paradigm that is normally used to detect age decrements in inferencing tends to emphasize the delayed component of comprehension. That is, the participant has to first read and remember large amounts of textual material before being asked to draw an inference (Hasher & Zacks. 1988). In the present study the format of the Making Inferences subtest presented a unique opportunity to examine inference comprehension in the condition of minimal memory load. The subtest provides inference statements to the participants in a multiple-choice format, together with the event chains. The clear failure of older adults on Making Inferences suggests that they were less able to establish a relationship between the explicit text information and possible inferences in a context where both the text and test information were presented simultaneously. Thus the results suggest that clear age-differences in inference resolution can in fact be observed in a condition where comprehension is examined immediately.

None of the cognitive mediators examined in the present study had an independent or strong contribution to performance on Making Inferences subtest. Working memory was not any more associated with this subtest than speed of processing or age per se. Furthermore, LTM which was expected to mediate the age differences on the Making Inferences also was not superior to other variables in independently accounting for individual differences on this subtest. Current literature suggests that inference processing may not necessarily heavily rely on processing resources. Similar to the present finding Masson & Miller (1983) and Dixon, Lefevre & Twilly (1988) found that working memory did not account for the unique variance in inference judgments. Hence the current findings suggest that other cognitive factors may be more important to performance on the Making Inferences subtest. One such factor might be the degree of contextual bias available during inference interpretation. Inferences are generally easier to draw when the information converges on a single possible interpretation, than when the contextual cues are compatible with several possible ideas (Lehman and Tompkins, 2001). As the Making Inferences subtest provides the participant with several possible inference resolution statements, reduced contextual bias may have differentially affected the abilities of the young and the old. Thus the necessity to edit away irrelevant or unnecessary concepts might have compromised the older adults' performance, while having minimal effect on the young adults' abilites. However, since the present research did not directly manipulate the type or degree of contextual bias, further studies are needed to explore these possibilities.

Performance of Elderly on the Recreating Sentences Subtest

Clear age differences were demonstrated on the Recreating Sentences subtest. Both of the older adult groups performed significantly worse than the young adult group. Furthermore, the old-old adults evidenced more impairment on this subtest than the young-old, confirming the presence of progressive age-related decline in production of syntactically complex sentences. This finding adds to the body of literature that has consistently demonstrated the presence of impairment in the ability of the elderly to construct grammatically complex sentences (Bromley, 1990, Kemper, 1987, Kynette & Kemper, 1986). Observations of participants' performance suggested that some of the elderly in the present sample had difficulty integrating and organizing the three words, which had to be incorporated into a sentence in a given context, and hence produced an intact sentence but with only one or two words. Other elders attempted to included all three words but produced sentences which were awkward, incomplete or semantically/pragmatically and syntactically inconsistent. For example, an older person's response to an item that asked to construct a sentence incorporating the words *neither, week, were*, in a supermarket context was: "*I have been here several weeks, but neither lettuces were not good price*".

In the present study the age decrements on the Recreating Sentences subtest was another example of unique mediation by working memory. Speed and long-term memory contributed to the subtest performance indirectly, through their effects on working memory. In this instance, however, age after controlling for speed, working memory and LTM still significantly predicted performance on the Recreating Sentences. The ability to construct grammatically complex sentences has already been linked to working memory function (Bock, 1982, Kemper, 1992, Kynette & Kemper, 1986, Kemper & Sumner, 2001). Harris (1994) also detected strong simple associations between working memory and this subtest. Bock (1982) has suggested that recreating sentences demands conceptual integration and synthesis of semantic, pragmatic and syntactic variables, and these processes of integration are likely to take place in working memory. From this perspective, the Recreating Sentence subtest placed heavy demands on working memory because of the requirement for the participants to construct sentences appropriate for specified contexts. To successfully accomplish the task the person had to be able to translate the information about the topic of the sentence, its perspective and so on via working memory into a form that can then activate the lexical and syntactic components of the system. The older adults' impairment in the operation of the working memory was partly responsible for the breakdown of sentence formation and production of syntactically incorrect or incomplete utterances. None of the mediator variables examined, however, were able to fully explain the age differences on the subtest. It is likely that language production is a complicated process with multiple determinants (Kemper, 1992). Ryan, Kwong See, Meneer & Trovato (1992) have presented a framework that emphasized not only cognitive but also social-cognitive factors (for e.g. reduced motivation to do well on the task, or even decreased physical activity (Hultsch et al, 1999) that could potentially influence the age differences. Future research will be needed to identify the age correlates in addition to cognitive factors that are important predictors of language production performance.

Performance of Elderly on the Figurative Language Subtest

The Figurative Language subtest was the least sensitive of the TLC-E subtests to age-related decline. Only the old-old group demonstrated some evidence of deterioration in performance. The present results are consistent with previous reports of persevered ability for metaphor interpretation until the late seventies (Light, et al., 1993). The Figurative Language subtest of evaluates the ability to interpret metaphoric expressions, and to match structurally related metaphors by shared meaning. Light et al (1993) postulated deficits in semantic processing as an explanation of decline in metaphor comprehension in the old-old elderly. The deficits are thought to arise due to the inability of the older elderly to form distinctive, contextually specific encodings of new information, and a tendency to encode the events in the "same old way" from one occasion to the next. This strategy may compromise the older adults' ability to understand distinctive, contextually new associations that are required by metaphor comprehension.

Out of all mediator variables examined in the present study, long-term memory was the only factor that successfully accounted for the individual differences on the subtest. The fact that working memory made only a small independent contribution to metaphoric expression interpretation is consistent with previous observations of weak working memory-figurative language associations (Light, et al., 1993, Harris, 1994). Studies that investigate metaphor comprehension generally indicate that the amount and quality of knowledge the person possesses concerning the topic and vehicle of the metaphor, plays a particularly important role in person's comprehension of the figurative expression (see Keil, 1986). In other words, the figurative language- long-term memory association possibly reflects age differences in the ability to gain access to the prior stored linguistic knowledge. This impairment was particularly evident when the older elderly performance was examined. In an easier condition of recognition the old-old adults were able to successfully identify the correct meaning of the metaphor. However, they experienced significant difficulties when un-cued recall of meaning was required.

In conclusion, older adults in this investigation presented difficulties on linguistic subtests relying on the interdependence of language and cognitive function. There was evidence of progressive age-related decline for the ability to understand ambiguous sentences, make inferences and recreate sentences. The ability to interpret metaphoric expressions was only comprised in the old-old. Working memory was the main factor that mediated age differences on ambiguous sentences and recreating sentences, while long-term memory was mostly responsible for individual differences on the Figurative Language task. None of the cognitive factors examined were able to successfully or uniquely account for performance on the Making Inferences subtest.

4.11 Decline in Cognitive Functioning in the Old-old Elderly

The current results highlighted the importance of examining the performance of both young-old and old-old subgroups. Most research studies on aging employ a wide range of ages beyond 60 or 65 and treat all the elderly as a homogeneous population. Such a strategy is clearly inadequate. In the present research, the old-old adults evidenced more profound decline on most tasks of working memory, speed of processing, inhibition and long-term memory. These results are in agreement with previous findings (Hultsch et al., 1999, Luszcz, 1992) that demonstrated significant changes in working memory, world knowledge, fluency and processing speed in old-old in comparison to the young-old. Indeed, longitudinal findings suggest that decrements at three year re-test intervals can be found on some tasks of processing skills (Hultsch, et al., 1992).

It should also be noted, however, that not all the measures of processing skills evinced progressive decline in the present study. For example, the old-old performed nearly at the same level as the young-old on Logical Memory, Daneman and Carpenter reading span, Spatial Span and Stroop interference. The examination of the pattern of scores suggested that the absence of any progressive changes on these measures was not due to the floor effects. One possible interpretation of these findings is that the abilities captured by these subtests may deteriorate at a different rate or that longer intervals (more than 10 years) may be required to observe age differences. The fact that story recall is relatively immune to progressive age decline replicates previous work (Colsher & Wallace, 1991, Hultsch, et al., 1999). These researchers suggested that story recall is less cognitively demanding than other recall tasks (i.e. Verbal Pairs) as it draws on the ability of the elderly to utilize contextual information to assist recall, a function that seems to be well preserved in the old-old (Kemper & Anagnopoulos, 1993). Merugo et al. (2000) have recently provided an interesting explanation for the attenuation of progressive deterioration on the reading span task in the old-old. The authors argued that the sharp difference between the young and the elderly on the span score reflects a decline in the capacity of the central executive system, whereas any further deterioration on the reading span score is a reflection of subsequent decline in capacity of the phonological loop, which itself is more resistant to early aging effects. In other words, it seems that different components of working memory might be affected during different stages of aging, with the central executive type functions declining at an earlier age.

Another issue is that it would be a mistake to assume that old age necessarily signifies a demise in cognitive functioning for all elderly individuals. The rate of cognitive aging may differ between individuals with physical and cognitive activity possibly serving as protective factors (see Hultsch, et al., 1999).

Examination of performance of the old-old elderly on the TLC-E clearly demonstrated that the discourse skills of this population decline profoundly in comparison to those of the young-old. This is a particularly interesting observation, especially if one considers that language functions are generally regarded as being more immune to age related loss (Schaie, 1983). The current results show that cohort differences in language functioning are more apparent on the tasks that require some degree of manipulation and organization of linguistic material (i.e. TLC-E tasks), while more crystallized verbal abilities (e.g. Vocabulary scores) do not suffer the same degree of decline. The obtained results are in concert with previous reports of progressive decline in language production (Kemper, et al., 2001, Heller & Dobbs, 1993). The important contribution of this study is that progressive decline was also documented on language comprehension tasks. Sparse information exists on the performance of the old-old on comprehension measures, apart from finding by Light et al. (1993) of figurative language impairment in the old-old. The old-old adults in the current study experienced more difficulty interpreting ambiguous sentences and making inferences in comparison to the young-old, which points to the presence of progressive decline in the ability to integrate, organize and interpret information from discourse. The TLC-E assesses language skills under natural discourse approximated constraints which permits to make tentative conclusions with regard to everyday language functioning in the elderly. As conversational discourse is normally repetitive and redundant the more profound impairments of the old-old may not be easily noticeable, but a decline would likely emerge when comprehension of propositionally dense and complex information (e.g. television news, technical text, newspaper articles) is required.

Overall, evidence of progressive age-related decline in some aspects of information processing and in higher language skills emphasizes the importance of examining cognitive change in several elderly cohorts. Such an approach provides a more accurate picture of the subtle cognitive deterioration that accompanies old age and reduces the risk of under or overestimating age-related changes.

4.12 Limitations of the Present Study

As with any research certain limitations can be identified in the present study. The main issues in this instance concern sample size, selection of measures of long-term memory, choice of language tests and choice of statistical techniques used.

The main limitation was sample size. Sixty-two people is the minimum number of participants required when normative investigations are undertaken. The possible negative effects the sample size may have on norms development were already discussed at the start of this chapter. It should be reiterated that although the present study did not recruit a large numbers of participants, the obtained sample was representative of the general population in terms of the proportional variety of educational background. Although the sample represented a wider community sample than is usually the case in neuropsychological investigations it was not a random sample. Limitations of the sample size might have also influenced the magnitude of the correlations observed. The performance of one or two outliers had the potential to substantially affect overall means. Nonetheless, scores on most measures demonstrated large variability and the correlations obtained were actually larger than reported in most published studies (see Salthouse, 1990, 1996 for comparison). Hence, the norms provided are worthwhile indicators but must be regarded as provisional. The conclusions based on mediational analyses seem to have not been adversely affected by sample size.

Another possible limitation of the present study was the choice of measures for long-term memory. The WMS-III subtests were selected to assess long-term memory performance. These subtests mainly assess learning and recall of new information. Measures that target more specifically retrieval and encoding functions (for example a version of an integration task that assesses knowledge access) may have been more suitable for the present purposes, given that current theories of long-term memory and working memory associations emphasize the interconnections between effective encoding/retrieval and working memory function (Ericsson & Delaney, 1999).

Furthermore, while the immediate memory scores were obtained, they did not feature in any further analysis. It would have been of interest to examine whether these measures, for example, are related to the functioning of Baddeley's episodic buffer.

In the field of cognitive aging two main research paradigms are normally utilized when language skills are evaluated: the on and off-line assessment of language comprehension. The present research evaluated language functioning using a unique task (TLC-E) that assesses language skills under natural, discourse approximated constrains. The TLC-E can be regarded as an off-line measure, but it differs from most standard off-line measures in terms of imposing minimal demands on retrieval of the actual test material, it does, however, still require active manipulation of linguistic information. It would have been of interest to examine how the pattern of results obtained using the TLC-E compares with that obtained when on-line measures of language processing are used as the on-line measures have the benefit of permitting the researcher to pinpoint the critical areas of the sentence that lead to difficulties in comprehension.

Other issues in the current study concern the statistical procedures used. Because simple correlational data is open to alternative causal interpretations, the path analysis technique is thought to provide a better estimate of causal relationships. However, path models should not be considered definitive because they can vary in the degree to which they represent or fit the data and alternative models often provide equally good fits. Hence, the path models can only be informative about the presence or absence of relationships. The path analysis results can also be affected by the degree of multicolinearity between the variables. Multicolinearity is a common problem in many correlational analyses where the two predictor variables are highly intercorrelated. In this case the certainty of the decision of which one of the two measures is a better predictor is considerably reduced. Interestingly, overviews of the relevant literature (e.g. Salthouse, 1991) indicate that researchers simply avoid addressing the problem of multicolinearity. They even often fail to report the simple correlations between the predictor variables.` Hence, the present study cannot ascertain whether the degree of multicolineraity observed here was more or less than that obtained elsewhere. Nonetheless the key pattern of findings reported here (namely the direct mediation by working memory of the age-language test associations) was replicated when the analysis was conducted using unstandardized residuals, that partialled out the effects of speed from all the variables of interest.

Salthouse (1990, 1996) often uses the hierarchical regression approach to ascertain the amount of age-related variance in cognition, that remains after controlling for the variance associated with speed and working memory. The outcome of this approach largely depends on the order in which the variables are entered. The variable that is entered first (which is always speed in Salthouse's investigations) absorbs most of the variance, with little being left for the next variable (i.e. working memory). In the current study when hierarchical regression analyses was conducted with speed as the first variable working memory was still associated with small but significant increase in R^2 (from .380 to .486) accounting for 10.6% of variance. However the reverse was not true when speed was added after the working memory, with increment in R^2 (from .475 to .486) amounting to only 1.1%. Thus, the hierarchical regression approach also supported the prominence of working memory as a mediating variable. Hence, it was clear that the main findings and conclusions in the current study were robust across various analytical techniques.

4.13 Summary of Contributions of the Present Study

The present study has been successful on a number of counts, with the last two listed below being the major contributions.

1) Clear impairments in naturally constrained discourse in older adults were observed.

2) A progressive deterioration in discourse skills has been demonstrated in the elderly, with a more profound impairment in the old-old adults in comparison to the young-old.

3) The presence of age-related decline on tasks that measure working memory capacity, speed of processing, resistance to interference and long-term memory was confirmed.

4) Preliminary evidence of superior performance of New Zealand young adults and young-old adults in comparison to the American normative population on the Wechsler Memory Scale-III Working memory and Auditory Memory subtests was obtained.

5) An indirect contribution of speed and long-term memory to discourse, through their effects on working memory, was found.

6) Evidence for the lack of association between inhibitory efficiency and discourse processing was obtained.

7) Some differential associations between various language skills assessed by the TLC-E subtests and their mediation by working memory, speed of processing, and long-term memory skills were identified.

8) Provisional (local) normative information for the Test of Language Competence-Expanded Edition for three age bands 20-34 years, 65-74 years and 75-89 years was obtained.

9) Clear evidence of unique mediation of age effects on discourse skills by working memory was demonstrated.

General Summary

In conclusion, the current study demonstrated that older adults experience progressive age-related decline in discourse and higher language skills. The study produced an informative set of data, including provisional norms on the performance of older adults, on the Test of Language Competence-Expanded Edition. The study also clearly demonstrated that the observed age-related decline in these higher language skills was mediated through age-related reductions in working memory, speed of processing and long-term memory, with working memory playing a crucial role in accounting for age differences. The findings provided by the present study will be useful for future evaluations of the contributions of working memory to language performance. The present study also highlighted the utility of the TLC-E as a promising measure for further evaluations of higher language function both in developmental studies (e.g. children, very old) and in clinical and experimental studies of various neurodegenerative disorders.

5. REFERENCES

Adams, B., Bell, L., Perfetti, C. (1995). A trading relationship between reading skill and domain knowledge in children's text comprehension. *Discourse-Processes*, 20, 307-323.

Albert, M., Wolfe, J. & Lafleche, G. (1990). Differences in abstraction ability with age. *Psychology and Aging*, 5, 94-100.

Allen, P., Madden, D., Weber, T. & Groth, K. (1993). Influence of age and processing stage on visual word recognition. *Psychology and Aging*, 8, 274-82.

Amrhein, P. & Theios, J. (1993). The time it takes elderly and young individuals to draw pictures and write words. *Psychology and Aging*, 8, 197-206.

Babcock, R. & Salthouse, T. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging*, 5, 421-8.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Physiology, 49* (1), 5-28.

Baddeley, A. (2000). Short-term and working memory. In E Tulving & F Craik (Eds.), *The Oxford handbook of memory*. (pp. 77-92). New York: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4, 417-423.

Baddeley, A. & Hitch, G. (1993). The recency effect: implicit learning with explicit retrieval? *Memory and Cognition*, 21, 146-55. Review.

Baddeley, A. & Logie, R. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 28-61). New York: Cambridge University Press.

Bashore, T., Osman, A. & Heffley, E. (1989). Mental slowing in elderly persons: a cognitive psychophysiological analysis. *Psychology and Aging*, 4, 235-44.

Bayles, K., Tomoeda, C. & Boone, D. (1985). A view of age-related changes in language function. *Developmental Neuropsychology*, 1, 231-264.

Beck, A., Steer, R. & Brown, G. (1996). *The Beck Depression Inventory-Second Edition.* New York: Harcourt Brace.

Belmore, S. (1981). Age-related changes in processing explicit and implicit language. *Journal of Gerontology*, 36, 316-322.

Bengston, V. & Schaie, K. (1989). *The course of later life: Research and Reflections.* New York: Sprinder Publishing Company.

Bock, J. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formation. *Psychological Review, 89*, 1-47.

Bowles, N. & Poon, L. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology, 40*, 71-77.

Bromley, D. (1991). Aspects of written language production over adult life. *Psychology and Aging*, 6, 296-308.

Burke, D. & Harrold, R. (1988). Automatic and effortful semantic processes in old age: experimental and naturalistic approaches. In L. Light & D. Burke (Eds.), *Language, Memory and Aging*. (pp. 100-116). New York: Cambridge University Press.

Burke, D. & Mackay, D. (1997). Memory, language, and ageing. *Philosohical Transcripts of the Royal Society of London, 352*, 1845-1856.

Burke, D. & Yee, P. (1984). Semantic priming during sentence processing by young and older adults. *Developmental Psychology, 20,* 903-910.

Cantor, J. & Engle , R. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory and Cognition, 5*, 1101-1114.

Caplan, D. & Waters, G. (1996). Syntactic processing in sentence comprehension under dual task conditions in aphasic patients. *Language and Cognitive Processes, 11,* 525-551.

Caplan, D & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77-126.

Carpenter, P., Miyake, A. & Just, M. (1994). Working memory constraints in comprehension. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego, CA: Academic Press.

Clegg, F. & Warrington, E. (2001). Psychometric testing of older adults: Provisional normative data for some commonly used tests.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology*, *11*, 412-429.

Cohen, G. (1981). Inferential reasoning in old age. *Cognition, 9*, 59-72.

Cohen, G., Dustman, R. & Bradford, D. (1984). Age-related decrements in Stroop color test performance. *Journal of Clinical Psychology, 40*, 1244-1250.

Colsher, P. & Wallace, R. (1991). Longitudinal application of cognitive function measures in a defined population of community-dwelling elders. *Annals of Epidimiology, 1*, 215-230.

Connelly, S., Hasher, L., Zacks, R. (1991). Age and reading: The impact of distraction. *Psychology and Aging, 6*, 533-541.

Covey, H. (1992). The definitions of the beginning of old age in history. *International Journal of Aging and Human Development, 34*, 325-337.

Cowan, N. (1999). An Embedded-Processes Model of working memory. In Miyake, A. (Ed); Shah, P. (Ed); et-al. *Models of working memory: Mechanisms of active maintenance and executive control.* (pp.62-101). New York: Cambridge University Press.

Craik, F. (1990). Changes in memory with normal aging: a functional view. *Advances in Neurology, 51*, 201-5. Review.

Daneman, M. & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Daneman, M. & Green, I. (1986). Individual differences in comprehending and producing words in context. *Memory and Language, 25*, 1-18.

Daneman, M. & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychometric Bulletin and Review, 13,* 422-433.

Dixon, P., LeFevre, J. & Twilley, L. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80,* 465-472.

Engle, R. (1996). Working memory and retrieval: An inhibition-resource approach. In J. Richardson, R. Engle, L. Hasher, R. Logie, E. Stoltzfus, R. Zacks (Eds.). *Working memory and human cognition*. New York: Oxford University Press

Engel, R., Cantor, J. & Carullo, J. (1992). Individual differences in working memory and comprehension: A test of four hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 972-992

Engle, R., Kane, M. & Tuholski, S. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 102-134). New York: Cambridge University Press.

Ericsson, K. & Delaney, P. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 257-297). New York: Cambridge University Press.

Folstein, M., Folstein, S & McHugh, P. (1975). Mini Mental State: A practical method of grading the cognitive state of patients for the clinicians. *Journal of Psychiatric Research, 12*, 189-198.

Foos, P. & Wright, L. (1992). Adult age differences in the storage of information in working memory. *Experimental Aging Research, 18*, 51-57.

Fraser, S., Glass, J. & Leathem, J. (1999). Everyday memory in an elderly New Zealand population: Performance on the Rivermead Behavioural Memory Test. *New Zealand Journal of Psychology, 28,* 118-123.

Gamboz, N., Russo, R., & Fox, E. (2000). Target selection difficulty, negative priming, and aging. *Psychology and Aging, 15*, 542-550.

Gibbs, R. (1994).Figurative thought and figurative language. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego: Academic Press.

Gold, J.; Carpenter, C., Randolph, C., Goldberg, T. & Weinberger, D. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry, 54,* 159-65.

Golden, C. (1978). *Stroop Color and Word Test.* New York; Stoeling and Co.

Grady, C. & Craik, F. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology, 10*, 224-231.

Grant, J. & Dagenbach, D. (2000). Further considerations regarding inhibitory processes, working memory, and negative priming. *American Journal of Psychology, 113,* 69-94.

Harris, R. (1994). *Discourse and non-discourse language and working memory in Alzheimer-type dementia*. Master of Arts in Psychology Thesis: University of Canterbury, New Zealand.

Hartley, J. (1993). Aging and prose memory: Tests of resource deficit hypothesis. *Psychology and Aging, 8*, 538-551.

Harvey, J. & Siegert, R. (1999). Normative data for New Zealand elders on the Controlled Oral Word Association Test, Graded Naming Test, and the Recognition Memory Test. *New Zealand Journal of Psychology, 28*, 124-132.

Hasher, L. & Zacks, R. (1988). Working memory, comprehension, and aging: a review and new view. *The Psychology of Learning and Motivation, 22,* 193-225.

Hasher, L., Zacks, R. & May, C. (1997). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application.* Cambridge, MA: MIT Press.

Heine, M., Ober, B. & Shenaut, G. (1999). Naturally occuring and experimentally induced tip-of-the-tongue experiences in three aldut age groups. *Psychology and Aging, 14,* 445-457.

Heller, R. & Dobbs, A. (1993). Age differences in word finding in discourse and nondiscourse situations. *Psychology and Aging, 8,* 443-450.

Hertzog, C., Raskind, C. & Canon, C. (1986). Age-related slowing in semantic information processing speed: an individual differences analysis. *Journal of Gerontology, 41*, 500-502.

Hopkins, K., Kellas, G. & Paul, S. (1995). Scope of word meaning activation during sentence processing by young and older adults. *Experimental Aging Research, 21*, 123-42.

Horn, J. (1982). The theory of fluid and cirstallized intelligencein relation to concepts of cognitive psychology and aging in adulthood. In f. Craik, & S. Trehub (eds.), *Aging and cognitive processes*. New York: Plenum Press.

Houx, P., Jolles, J. & Vreeling, F. (1993). Stroop interference: Aging effects assessed with Stroop color-word test. *Experimental Aging Research, 19*, 209-224.

Hultsch, D. & Dixon, R. (1983). The role of pre-experimental knowledge in text processing in adulthood. *Experimental Aging Research, 9,* 17-22.

Hultsch, D., Hertzog, C., Small, B. & Dixon, R. (1999). Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging? *Psychology and Aging, 14*, 245-63.

Just, M. & Carpenter, P. (1980). A theory of reading : From eye fixation to reading comprehension. *Psychological Review, 99,* 122-149.

Just, M. & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kausler, D. (1994). *Learning and memory in normal aging.* San Diego: Academic Press, Inc.

Kausler, D. & Wiley, J. (1991). Effects of short-term retrieval on adult age differences in long-term recall of actions. *Psychology and Aging, 6,* 661-5.

Keil, F. (1986). Conceptual domains and the acquisition of metaphor. *Cognitive Development, 1*, 73-96.

Kellas, G., Paul, S. & Vu, H. (1995). Aging and language performance: From isolated words to multiple sentence contexts. In Ph. Allen & Th. Bashore (Eds.), *Age diffirences in word and language processing*, NY, Elsevier Science.

Kemper, S. (1986). Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics, 7*, 277-287.

Kemper, S. (1987). Syntactic complexity and the recall of prose by middle-ages and elderly adults. *Experimental Aging Research, 13,* 47-52.

Kemper, S. (1988). Geriatric psyholinguistics: Syntactic limitations of oral and written language. In L. Light & D. Burke (Eds.), *Language, memory and aging.* New York, Cambridge University Press.

Kemper, S. (1990). Adult's diaries: Changes made to written narratives across the life-span. *Discourse Processes, 13,* 207-223.

Kemper, S. (1992). Language and aging. In F.I.M.Craik & T.A. Slathouse (Eds.), *The handbook of aging and cognition*. Hillsdale, NJ: Erlbaum.

Kemper, S. (1997). Metalinguistic judgements in normal aging and Alzheimer's disease. *Journal of Gerontology: Psychological Sciences, 52*, 147-155.

Kemper, S. & Anagnopoulos, C. (1993). Adult use of discourse constraints on syntactic processing. In J Cerella & J Rybash (Eds.), *Adult information processing: Limits on loss*. New York: Academic Press.

Kemper, S., Greiner, L., Marquis, J., Prenovost, K. & Mitzner, T. (2001). Language declines across the life span: Findings from the nun study. *Psychology and Aging, 16*, 227-239.

Kemper, S. & Kemtes, K. (2000). Aging and message production and comprehension. In D. Park & N. Schwarz (Eds.), *Aging and cognition: a primer*. Philadelphia: Psychology Press.

Kemper, S. & Kliegl, R. (1999). Concluding observations. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, memory, and grammer* (pp. 299-307). New York: Kluwer Academic.

Kemper, S. & Mitzner, T. (2001). Language production and comprehension. In J. Birren & K. Warner Schaie (Eds.), *Handbook of the Psychology of aging*. San Diego: Academic Press.

Kemper, S. & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16,* 312-322.

Kemtes, K. & Kemper, S. (1996). *Younger and older adults' working memory and on-line processing of syntactically ambiguous sentences.* Poster presented at the Cognitive Aging Conference, Atlanta, GA.

Kemtes, K. & Kemper, S. (1997). Younger and older adults' on-line processing of syntactically ambiguous sentences. *Psychology and aging, 12* (2), 362-371.

Kemtes, K. & Kemper, S. (1999). Aging and the resolution of quantifier scope ambiguities. *Journal of Gerontology: Psychological Sciences, 54B*, 350-360.

Kemtes, K. & Kemper, S. (2001). Cognitive construct measurement in small samples of younger and older adults: an example of verbal working memory. *Experimental Aging Research, 27*, 167-80.

Kieley, J. & Hartley, A. (1997). Age-related equivalence of identity suppression in the Stroop color-word task. *Psychology and Aging, 12*, 22-9.

Kirasic, K., Allen, G., Dobson, S. & Binder, K. (1996). Aging, cognitive resourcers and declarative learning. *Psychology and Aging, 11*, 658-670.

King, J. & Just, M. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580-602.

Kramer, A., Humphrey, D., Larish, J, Logan, G. & Strayer, D. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processes in attention. *Psychology and Aging, 4,* 491-512.

Kwong See, S. & Ryan, E. (1995). Cognitive mediation of adult age differences in language performance. *Psychology and Aging, 10,* 458-468.

Kynette, D. & Kemper, S. (1986). Aging and the loss of grammatical forms: A cross-sectional study of language performance. *Language and Communication*, 6, 65-72.

Lehman, B. & Tompkins, C. (2001). Predictive inferencing in adults with right hemisphere brain damage. *Journal of Speech, Language, and Hearing Research*, *44,* 639-654.

Lethlean, J. & Murdoch, B. (1997). Performance of subjects with multiple sclerosis on tests of high-level language. *Aphasiology, 11,* 1, 39-57.

Lewis, F., Lapointe, L., Murdoch, B & Chenery, H. (1998). Language impairment in Parkinson's disease. *Aphasiology, 12*, 193-206.

Light, L. (1988). Language and aging: Competence versus performance. In J. Birren & V. Bengston (Eds.), *Emergent theories of aging*. New York; Springer.

Light, L. (1990). Interactions between memory and language in old age. In J. Birren & K. Shaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 275-290), San Diego, CA: Academic Press.

Light, L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Revue of Psychology*, *42*, 333-375.

Light, L. & Albertson, S. (1988). Comprehension of pragmatic implications in young and older adults. In L. Light & D. Burke (Eds.), *Language, memory and aging*. New York: Cambridge University Press.

Light, L. & Anderson, P. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology*, 40, 737-747.

Light, L. & Burke, D. (1988). Patterns of language and memory in old age. In L. Light & D. Burke (Eds.), *Language, memory and aging*. New York: Cambridge University Press.

Light, L. & Capps, J. (1986). Comprehension of pronouns in younger and older adults. *Developmental Psychology, 22,* 580-585.

Light, L., Owens, S., Mahoney, P. & LaVoie, D. (1993). Comprehension of metahpors by young and older adluts. In J Cerella & J Rybash (Eds.), *Adult information processing: Limits on loss*. New York: Academic Press.

Light, L., Zelinski, E. & Moore, M. (1982). Adult age differences in reasoning from new information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8,* 435-447.

Luszcz, M. (1992). Predictors of memory in young-old and old-old adults. *International Journal of Behavioural development, 15,* 147-166.

McCarty, S., Siegler, I. & Logue, P. (1982). Cross-sectional and longitudinal patterns of three Wechsler Memory Scale Subtests. *Journal of Gerontology, 37,* 169-75.

MacDonald, M., Just, M. & Carpenter, P. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56-98.

McGinnis, D. & Zelinski, M. (2000). Understanding unfamiliar words: The influence of processing resources, vocabulary knowledge, and age. *Psychology and Aging, 15,* 335-350.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203.

Madden, D. (1986). Adult age differences in visual word recognition: Semantic encoding and episodic retention. *Experimental Aging Research, 12*, 71-77.

Madden, D. (1989). Visual word identification and age-related slowing. *Cognitive Development, 4*, 1-29.

Malec, J., Invik, R. & Smith, G. (1993). In R. Parks, R. Zec & R. Wilson (Eds.). *Neuropsychology of Alzheimer's disease and other dementias*. New York: Oxford University Press.

Marmurek, H. (1989). Familiarity effects and word unitization in visual comparison tasks. *Memory and Cognition,17*, 483-489.

Masson, M. & Miller, J. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology, 75*, 314-318.

May, C., Zacks, R., Hasher, L. & Multhaup, K. (1999). Inhibition in the processing of garden-path sentences. *Psychology and Aging, 14,* 304-13.

Mayr, U. & Kliegl, R. (2000). Complex semantic processing in old age: Does it stay or does it go? *Psychology and Aging, 15*, 29-43.

Meguro, Y., Fujii, T., Yamadori, A., Tsukiura, T., Suzuki, K., Okuda, J. & Osaka, M. (2000). The nature of age-related decline on the reading span task. *Journal of Clinical and Experimental Neuropsychology, 22*, 391-8.

Mitchell, D. (1994). Sentence parsing. In M Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego: Academic Press.

Miyake, A., Carpenter, P. & Just, M. (1994). A capacity approach to syntactic comprehension: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11*, 671-717.

Miyake, A., Carpenter, P. & Just, M. (1995). Reduced resourcers and specific impairements in normal and ahpasic sentence comprehension. *Cognitive Neuropsychology, 12,* 651-679.

Miyake, A. & Shah, P. (1999). *Modles of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge Press.

Nelson, H. (1991). *National Adult Reading Test-Second Edition.* England: NFER- Nelson Publishing Company.

Nettelbeck, T. & Rabbitt, P. (1992). Aging, cognitive performance and mental speed. *Intelligence, 16*, 189-205.

Norman, S., Kemper, S. & Kynette, D. (1992). Adult's reading comprehension: Effects of syntactic complexity and working memory. *Journal of Gerontology: Psychological Sciences, 47*, 258-265.

Norman, S., Kemper, S., Kynette, D., Cheung, H. & Anagnopoulos, C. (1991). Syntactic complexity and adults' running memory span. *Journal of Gerontology*, *46*, 346-351.

Obler, L. & Albert, M. (1985). Language skills across adulthood. In J. Birren & K. Schaie (Eds.), *Handbook of the Psychology of Aging.* (pp. 463-473). New York: Van Nostrand Reinhold. 2nd ed.

Obler, L., Fein, D., Nicholas, M. & Albert, M. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics, 12,* 433-452.

Onifer, W. Swinney, D. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition*, *9*, 225-236.

Palladino, P. & De Beni, R. (1999). Working memory in aging: maintenance and suppression. *Aging, 11*, 301-306.

Pratt, M . & Robins, S. (1991). That's the way it was: Age differences in the structure and quality of adults' personal narratives. *Discourse Processes, 14*, 73-85.

Rogers, W. & Fisk, A. (1991). Age-related differences in the maintenance and modification of automatic processes: arithmetic Stroop interference. *Human Factors*, *33,* 45-56.

Rosen, V. & Engle, R. (1994). *Working memory capacity and retrieval from long-term memory.* Unpublished manuscript.

Rosen, V. & Engle, R. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology,126*, 211-27.

Ryan, E., Kwong-See, S., Meneer, W. & Trovato, D. (1992). Age-based perceptions of language performance among young and older adults. *Communication Research, 19*, 423-443.

Salthouse, T. (1980). Age and memory: Strategies for localising the loss. In L. Poon, J. Fozard, L. Cermak, D. Arenburg & L. Thompson (Eds.), *New directions in memory and aging: Proceedings of the George A. Talland memorial conference.* Hillsdale, NJ: Erlbaum.

Salthouse, T. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale: NJ:Erlbaum.

Salthouse, T. (1992). Shifting levels of analysis in the investigation of cognitive aging. *Human Development, 35*, 321-342.

Salthouse, T. (1993). Speed mediation of adult age differences in cognition. *Developmental Psychology, 29,* 722-738.

Salthouse, T. (1994). The aging of working memory. *Neuropsychology, 8*, 535-543.

Salthouse, T. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Salthouse, T. & Babcock, R. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27*, 763-776.

Salthouse, T., Babcock, R., Shovronek, E., Mitchell, D. & Palmon, R. (1990). Age and experiience effects in spatial visualisation. *Developmental Psychology, 26,* 128-136.

Salthouse, T., Kausler, D. & Saults, J. (1988). Utilization of path-analytic procedures to investigate the role of processing resources in cognitive aging. *Psychology and Aging, 3,* 158-66.

Salthouse, T. & Meinz, E. (1995). Aging, inhibition, working memory, and speed. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 50B*, 297-306.

Schaie, K. (1983). The Seattle longitudinal study: a 21 year exploration of psychometric intelligence. In K. Schaie (Ed.). *Longitudinal studies of adult psychological development.* New York: Guilford Press.

Schaie, K. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging, 4*, 443-453.

Schaie, K. & Hofer, S. (2001). Longitudinal studies in aging research. In Birren, J & Warner Schaie, K (Eds.), *Handbook of the psychology of aging.* San Diego: Academic Press.

Schumacker, R. & Lomax, R. (1996). *A beginner's guide to structural equation modeling.* New Jersey: Lawrence Erlbaum.

Shah, P. & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4-27.

Shah, P. & Miyake, A. (1999). Models of Working Memory: An Introduction. In P. Shah & A. Miyake (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 1-27). New York: Cambridge University Press.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge: Cambridge University Press.

Siegler, I., McCarty, S. & Logue, P. (1982). Wechsler Memory Scale Scores, selective attrition, and distance from death. *Journal of Gerontology, 37,* 176-81.

Simpson, G. (1994). Context and the processing of ambiguous words. In M Gernsbacher (Ed.), *Handbook of Psycholinguistics.* San Diego: Academic Press.

Singer, M. & Ritchot, F. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition, 24,* 733-743.

Sliwinski, M. & Buschke, H. (1999). Cross-sectional and longitudinal relationships among age, cognition, and processing speed. *Psychology and Aging, 14*, 18-33.

Smith, G., Malec, J. & Ivnik, R. (1992). Validity of the construct of nonverbal memory: a factor-analytic study in a normal elderly sample. *Journal of Clinical Experimental Neuropsychology, 14,* 211-21.

Snowdon, D. (1997). Aging and Alzheimer's disease: Lessons from the Nun Study. *Gerontologist, 37*, 150-156.

Statistics New Zealand (1996). *Census.* New Zealand

Stine, E. (1990). On-line processing of written text by younger and older adults. *Psychology and Aging, 5*, 68-78.

Stine, E. (1995). Age and the distribution of resources in working memory. In P. Allen & T. Bashore (Eds.). *Age differences in word and language processing.* New York: Elsevier Science.

Stine, E., Cheung, H. & Henderson, D. (1995). Adult age differences in the on-line processing of new concepts in discourse. *Aging and Cognition, 2,* 1-18.

Stine, E., & Hindman, J. (1994). Age differences in reading time allocation for propositionally dense sentences. *Aging and Cognition, 1,* 2-16.

Stine, E. & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging, 2*, 272-279.

Stine, E. & Wingfield, A. (1988). Memorability functions as an indicator of qualitative age differences in text recall. *Psychology and Aging, 3*, 179-83.

Stine-Morrow, E., Loveless, M. & Soederberg, L. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology and Aging, 11*, 475-486.

Stuss, D., Shallice, T., Alexander, M. & Picton, T. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of New York Academy of Science, 769*, 191-211.

Till, R. (1985). Verbatim and inferential memory in young and elderly adults. *Journal of Gerontology, 40*, 316-23.

Till, R. & Walsh, D. (1980). Encoding and retrieval factors in adult memory for implicational sentences. *Journal of Verbal Learning and Verbal Behavior, 19*, 1-16.

Troyer, A., Moscovitch, M. & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology, 11,* 1, 138-146.

Tun, P., Wingfield, A. & Stine, E. (1991). Speech-processing capacity in younger and older adults: A dual task study. *Psychology and Aging, 6*, 3-9.

Tun, P., Wingfield, A., Stine, E. & Mecsas, C. (1992). Rapid speech processing and divided attention: Processing rate versus processing resources as an explanation of age effects. *Psychology and Aging, 7*, 546-550.

Tun, P. & Wingfield, A. (1993). Is speech special? Perception and recall of spoken language in complex environments. In J. Cerella & J. Rybash (Eds.), *Adult information processing: Limits on loss.* (pp. 425-457). San Diego: Academic Press, Inc.

Ulatowska, H., Cannito, M., Hayashi, M. & Fleming, S. (1986). Language abilities in the elderly. In H. Ulatowska (Ed.), *The aging brain: Communication in the elderly.* San Diego: College-Hill.

Ulatowska, H., Chapman, S., Highley, A. & Prince, J. (1998). Discourse in healthy old-elderly adults: a longitudinal study. *Aphasiology, 12*, 619-633.

Van der Linden, M., Beerten, A. & Pesenti, M. (1998). Age-related differences in random generation. *Brain and Cognition, 12,* 313-325.

Van der Linden, M., Bredart, S. & Beerten, A. (1994). Age-related differences in updating working memory. *British Journal of Psychology, 85*, 145-152.

Van der Linden, M., Hupet, M., Feyereisen, P., Schelstraete, M., Bestgun, Y., Bruyer, R., Lories, G., El Ahmadi, A. & Seron, X. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology, and Cognition, 6,* 32-55.

Verhaeghen, P., Marcoen, A. & Goossens, L. (1993). Facts and fiction about memory aging: a quantitative intergration of research findings. *Journal of Gerontology, 48,* 157-171.

Waters, G. & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging, 16*, 128-144.

Wechsler, D. (1975). Intelligence defined and undefined: A realistic appraisal. *American Psychologist, 30*, 135-139.

Wechsler, D. (1997). *Wechsler Memory Scale-Third Edition.* San Antonio: Harcourt Brace and Company.

Wechsler, D. (1999). Wechsler Abbreviated Adult Intelligence Scale. San Antonio: Harcourt Brace and Company.

West, R. (1999). Age differences in lapses of intention in the Stroop task. *Journal of Gerontology, 54B*, 34-43.

Wiig, E. & Secord, W. (1989). *Test of Language Competence.* San Antonio, Texas: Psychological Corporation.

Wingfield, A., Stine, E., Lahar, C. & Aberdeen, J. (1988). Does the capacity of working memory change with age? *Experimental Aging Research, 13,* 103-107.

Wingfield, A. & Stine-Morrow, E. (2000). Language and Speech. In F. Craik & T. Salthouse (Eds.), *Handbook of aging and cognition (2nd ed., pp. 359-416).* Mahwah, NJ: Erlbaum associates.

Zelinski, E. & Miura, S. (1990). Anaphor comprehension in younger and older adults. *International Journal of Aging and Human Development, 31*, 111-34.

6. APPENDIX

6.1

Advertisement

VOLUNTEERS WANTED

Participants wanted for a study looking at
language, memory and attention in younger and older people.

If you are aged between 20-34 years
or you are aged between 65-89 years,
and English is your first language, then we need you!

$30 is provided to cover travel expenses

For more information contact Elena
Phone: 381-3354
e-mail: jpm61@student.canterbury.ac.nz

6.2

INFORMATION ABOUT THE PROJECT

You are invited to participate as a subject in the research project **"Language competence and working memory in older adults" (Contact: Elena Loukavenko 381-3354)**

The aim of this project is to examine whether people's language skills change with age. We know that as a person grows older his or her memory may show some change, but less is known about language skills or the relationship between memory and language. The present research will attempt to answer these questions.

In this project, you will be asked to do a short series of tasks that assess certain language skills, memory and attention. For example, memory and attention tasks involve memorizing a short story or repeating some numbers. Examples of the language tasks include making up simple sentences by looking at the picture or making an inference from two statements. Your responses on some of those tests will be audiotaped to enable us to score them accurately. You will also be asked to provide some general information, which covers the kind of background information that researchers need for scientific studies. That information includes standard checklists for concentration, orientation, word knowledge, reading and writing. One checklist includes responses that could be interpreted as indicating the presence or absence of depression (if your score on this checklist does indicate the possibility of detecting depression, we will of course inform you and advise you to contact your GP for further evaluation).

To prevent the chance of fatigue affecting the performance of any individual, you would be asked to attend two separate testing sessions. Each session will take up to 2.5 hours. These sessions will be scheduled on dates and times most convenient for you. You will also have an opportunity to have short breaks during the sessions. The research will be conducted at Canterbury University. We will provide thirty dollars to cover your travel expenses.

At any time during the project you have the right to withdraw your participation and any information provided. It is not anticipated that participation in the study will involve any risk to you.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation and your identity will not be made public. To ensure anonymity and confidentiality, all the identifiable information will be securely stored in a locked safe in the Psychology Department Canterbury University premises. An electronic version of the information, using a coded ID number will be stored on the researcher's personal computer access to which is protected by password. The information collected may be used for future research projects, for which a separate approval from the Human Ethics Committee will be obtained, and your anonymity and confidentiality will continue to be preserved.

This project is being undertaken for a Masters of Science degree by Elena Loukavenko under the supervision of Dr. John Dalrymple-Alford, who can be contacted at 364-2998 or 364-2994 (or email psyc338@psyc.canterbury.ac.nz). He will be pleased to discuss any concerns you may have about participation in the project.

To arrange or change times for participation in the study please contact Elena Loukavenko 381-3354.
Otherwise, please e-mail Elena at: jpm61@student.canterbury.ac.nz

The project has been reviewed and approved by the University of Canterbury Human Ethics Committee.

6.3

CONSENT FORM

Study:
Language competence and working memory in older adults

I have read and understood the description of the above-named project. On this basis I agree to participate as a subject in the project. I also agree that some of my responses on the tests will be audiotaped with an understanding that after the scoring is completed the recording will be destroyed. I consent to publication of the results of the project. I also agree that the information collected in this project may be used in future research on the understanding that separate approval will be obtained from the Human Ethics Committee at that time. I understand that my anonymity will be preserved both in the publication of the results and any future use of the information collected. I understand also that I may at any time withdraw from the project, including withdrawal of any information I have provided.

Signed..

Date..

6.4 **ID number: _______**

Health and Information Questionnaire
Filled in by Experimenter

PARTICIPANT INFORMATION

1) Surname :________________________________

First/Given name:___________________________________

2) Address:____

Contact tel. Number (s):________________________________

3) Gender: Male Female

4) Date of birth: Day_____ Month________Year________

5) What is your first language (the language you and your family would speak at home)?_________________________

6) Years of education? (in years post age 10/11, that is <u>not</u> counting primary school)__________

7) Qualifications:

□ School qualification (For example: school certificate passes, sixth form qualification, higher school qualification, University Bursary Entrance Examination).

□ Vocational qualification (For example: trade certificate, technicians certificate, apprenticeships, national certificate, national diploma, advanced trade certificate bringing certificate, pre vocational certificate).

□ Higher qualification (For example: undergraduate diploma or certificate, New Zealand diploma or certificate, BA, BSc, MA, Ph.D., post-graduate diploma).

□ None of the above

8) Which day of the week and time you are most likely to be free to take part in the study?

9) General Health Questions.

The Following statement is read by the researcher prior to asking these questions:

"I am going to ask you some general questions on various health problems that people may experience. This information is strictly confidential; you are <u>*not*</u> *obliged to answer any of these questions or any particular question."*

1) Have you ever suffered in the past or presently suffering from any of the following conditions?

a) Moderate or severe Head injury/ Stroke/ or other neurological impairment (for example: Multiple Sclerosis, Parkinson's disease).

b) Major medical illness (for example: a history of severe migraine, major heart condition, diabetes requiring insulin injections).

c) Any significant psychiatric illness requiring hospitalization.

d) Major depression in last 6 months

e) Any learning disability

2) Do you currently take any medications that you think may affect your performance today? (this question to be repeated for each session).

3) Are you currently involved in any therapeutic trial (exclude if "yes")?

4) Please rate any use of the following:

a) Caffeine (coffee, tea, chocolate, caffeinated soft drinks)

<u>Per day</u>: None / Little (One cup or can)
Moderate (2 or 3)
Heavy (4 or more)

b) Alcohol <u>Daily average (ALAC guide: Male amount shown; halve this for women)</u>

<u>Per day:</u> None / Little (less than moderate daily average)
Moderate (1/3 to 1 spirit; 1-2.5 glass wine; or 2-5 beer)
Heavy (more than moderate daily average)

c) Recreational / mood altering drugs (based on average)

None / Little (once per month)
Moderate (once or twice per week)
Heavy (daily)

5) Do you have normal vision (or vision that is normal when corrected by glasses/contact lenses?)

6) Do you have normal hearing (or hearing that is normal when corrected with an aid?)

6.5

Statement Regarding the Beck Depression Inventory Score

Study on Language Competence and Working Memory in Older Adults

This study is designed to provide scientific research information. It is not a formal clinical evaluation. The Beck Depression Inventory Score (called the BDI score) is not our main interest, but scientific research usually adjusts other scores on some tasks to rule out any separate effects of the BDI score. The score can, however, be used by other clinical workers in combination with a clinical evaluation to assess symptoms of depression.

In your case, this score could indicate a [ ] level of depressive symptoms. Therefore, we ask that you please complete the following:

I understand that this score is only indicative of the possibility of depression, but I have been advised by the Researcher that I should in the first instance contact my GP for further evaluation, should I choose.

I agree / disagree (Participant to circle their preference and cross out word which they find inappropriate) that the Researcher may contact me as a follow-up reminder, AND

I agree / disagree (Participant to circle their preference and cross out word which they find inappropriate) that the Researcher may contact a friend or relative or my GP in confidence to provide me with further advice (Give name / contact here, if appropriate) . I understand that this contact would be made concerning my BDI score only, not any other score or information provided during this study.

Name and phone of contact – only if agreed by the participant:____________________

I, __________________________ (print full name) fully understand the above statement, as amended by me, and understand that I will be given a signed copy of this statement.

Signature of Participant_________________________Date____________________

Signature of witness (the researcher) ________________

Name of witness__________________________

Date:______________________________

6.6

Items and Instructions for the Reading Span Task

Instructions:

"In this task you will be presented with a series of unrelated sentences displayed on the cards. Whenever the sentence is presented to, you are to read it out loud. Some of the sentences make sense, and some of the sentences do not make sense. After you finish reading the sentence, I want you to say, "Yes" if the sentence makes sense and "No" if the sentence does not make sense. When you are deciding whether the sentence makes sense, keep in mind that I am not trying to trick you with hidden meanings or anything, so don't waste too much mental energy over analyzing the sentences. After you answer "Yes" or "No" I will turn the card over and show you the next sentence, again you will have to say "Yes" or "No" to indicate whether or not the sentence makes sense. Keep doing this until we get to a blank card. The blank card means that the trail is over, and you have to say back to me the last word in each of the sentences in the trial.

So here's an example: (Show practice item 1)
If possible, you are to say back the last words in the order in which they were presented. If you can't remember them in order, you can say them in any order, but you should not start with the last word first, unless it is the only one you can remember.

Your goal is to try to say back as many of the last words in the trial as possible.
We will be starting off with trials consisting of two sentences and will periodically increase the number of sentences per trial without any advance warning. That is, we will progress to three-sentence trials, then four sentence trials and so on.

The first couple of trials are for practice so you can get the hang of it.

Reading Span Items

Practice Items (all at 2 sentence length)

Set 1
The house quickly got dressed and went to work.
I took a knapsack from my shovel and began removing the earth.

Set 2
The lamp bucked and sent the horse tumbling to the ground.
The cop spent a good half-hour questioning his trusted friend.

Set 3
People are given by money at Christmas time.
She worked quickly and quietly while others were asleep.

Set 4
The sun had gone and the evening skies were tinted purple.
Opposite the chimney doorway was the yawning cabin mouth.

Set 5
A deafening cheer rose up from the kids watching the parade.
A blue-uniformed security guard moved quickly out of the dog.

Test Items

Level 2 sets

Set 1
It was a foggy day and everything was dripping wet.
The girl was awakened by the gusts of rain blowing against the house.

Set 2
The story started as a joke but soon got out of hand.
He quickly put the carrot in the ignition and started the car.

Set 3
The starving hamburger bit into the juicy man.
The hurricane left a path of destruction through the tiny town.

Level 3 sets

Set 1
The murky swamp slipped into the waters of the crocodile.
The castle sat nestled in the refrigerator above the tiny village.
It wasn't all her fault that her marriage was in trouble.

Set 2
When he reached the top of the heart his mountain was pounding.
The barn raged through the abandoned old fire.
With the frown of pain, the old ranger hung up his hat forever.

Set 3
The man fidgeted nervously, once again checking his watch.
Clouds of cigar smoke wafted into the open eraser.
Convictions for all offences increased from the turn of the century.

Level 4 sets

Set 1
They waited at the water's edge, the raft bobbing up and down.
I let the potato ring and ring, but still no answer.
The red wine looked like blood on the white carpet.
The children put on their closets and played in the snow.

Set 2
At some life, everyone ponders the meaning of point.

The bars roared and began banging on the ape of the cage.
Being sued for malpractice was the doctor's main concern.
The shampoo was vibrant with music, theatre, and dance.

Set 3
The class homework was done by everyone in the history.
Thick foliage surrounded him, and the air was heavy and still.
The deserted calendars rocked mournfully, driven by the tide.
The men were all killed during the training flight near the base.

Level 5 sets

Set 1
An eerie breeze suddenly chilled the warm, humid air.
As the ideas flowed, I jotted them down on some water.
The flask was dark, lit only by the occasional room of lightning.
He stepped back as the ghoul moved forward.
The robber bounded across the bridge and entered the dimly lit garage.

Set 2
Three of the pillows were dead and he was next.
My escape out of the telephone was blocked by a wire fence.
She turned around and sucked in a startled breath.
They ran until their lungs felt like they were going to burst.
The additional evidence helped the verdict to reach their jury.

Set 3
No one ever figured out what caused the crash to plane.
His eyes were bloodshot and his face was pale.
As a full-time student, he studied hard.
The tower raced across the sailboat to the finish line.
Somewhere in the deepening twilight, a loon sang its evening song.

Level 6 sets

Set 1
Trails are supposed to stay on the hikers, but they usually don't.
He stormed out without giving me so much as a backward glance.
The paperclip was flaked white and red with sunburn.
Returning with an eagle, a branch breaks to land at its nest.
A television droned from the dark interior of the apartment.
They talked about what the world would be like after the war.

Set 2
His mouth was twisted into an inhuman smile.
The closet doors were wrenched open.
A welt was forming on his bottle where the forehead made contact.
I'd been naïve to think he would fall into my trap.
The piercing yellow eyes glowed hauntingly in the mist.

The beach was filtering the moonlight from outside.

Set 3

These operations are only done as a last resort.
The first impression is often a lasting one.
The throat tightening around her arm turned her scream into a croak.
The soap hovered over the elephant, waiting to attack.
They watched in silence as a new carpet dipped behind the horizon.
The rumbling of the distance faded into a feather.

6.7
Instructions and Category Exemplars for the Semantic Fluency Task

Instructions:
Administration: **Time 5 minutes**

"For this next task, when I say "GO" I want you to tell me as many names of animals as you can remember. You will have to keep going until I tell you to stop. And please, do not repeat names you have already said. For this task I am going to use a tape recorder, so I can write down exactly what you've said later on. Any questions? Ready? GO!"

Category Exemplars

Note that the list is only indicative and includes most commonly produced exemplars.

Living Environment

Africa: aardvark, antelope, buffalo, camel, chameleon, cheetah, chimpanzee, cobra, eland, elephant, gazelle, giraffe, gnu, gorilla, hippopotamus, hyena, impala, leopard, lion, manatee, mongoose, monkey, ostrich, panther, rhinoceros, tiger, wildebeest, warthog, zebra.

Australia: alligator, budgerigar, cockatoo, cockateel, dingo, emu, kangaroo, opossum, parakeet, platypus, Tasmanian devil, wallaby, wombat.

***Arctic/Far North*:** auk, caribou, musk ox, penguin, polar bear, reindeer, seal, whale.

Farm: chicken, cow, donkey, goat, hen, horse, mule, ox, pig, rooster, romni, sheep, turkey

New Zealand: bellbird, fantail, gecko, kiwi, kea, kakapo, moa, morpok, possum, red-back spider, tar, tui, tuatara, weka, waxeye, wood pigeon, yellow eyed penguin.

North America: badger, bear, beaver, bobcat, caribou, chipmunk, cougar, deer, elk, fox, moose, mountain lion, puma, rabbit, raccoon, skunk, squirrel, wolf.

Water: alligator, auk, beaver, crocodile, dolphin, eel, fish, frog, lobster, manatee, muskrat, newt, octopus, otter, oyster, penguin, platypus, salmon, salamander, sea elephant, sea lion, seal, shark, toad, trout, turtle, whale

Human use
Beasts of burden: camel, donkey, horse, llama, ox

Fur: beaver, chinchilla, fox, mink, rabbit

Pets: budgie, bulldog, canary, cat, corgi, Dalmatian, dog, gerbil, golden retriever, guinea pig, hamster, parrot, terrier, rabbit

Zoological Categories

Bird: albatross, budgie, bellbird, blackbird, bird, crow, crane, condor, cockatoo, dove, eagle, emu, fantail, finch, flamingo, humming bird, ibex, fantail, finch, kiwi, kea, moa, morpok, ostrich, owl, parrot, pelican, penguin, robin, sparrow, swan, seagull, tit, tui, quail, toucan, weka, woodpecker, waxeye, wood pigeon

Bovine: bison, buffalo, cow, musk ox, yak

Canine: coyote, bulldog, Dalmatian, dog, fox, fox terrier, hyena, jackal, spaniel, wolf.

Deer: antelope, caribou, chamuar eland, elk, gazelle, gnu, impala, moose, reindeer, wildebeest, tar

Feline: bobcat, cat, cheetah, cougar, jaguar, leopard, lion, lynx, mountain lion, ocelot, panther, Persian, puma, tabby, tiger.

Fish: bass, brown trout, blue cod, guppy, groper, eel, flounder, harpuka, herring, ling, piranha, salmon, sardines, trout, tuna, taraki.

Insect: ant, beetle, butterfly, cockroach, flea, fly, ladybird, mosquito, moth, praying mantis, spider, wasp.
Insectivores: aardvark, anteater, hedgehog, mole, shrew

Primate: ape, baboon, chimpanzee, gibbon, gorilla, human, lemur, marmoset, monkey, orangutan, shrew

Rabbit: coney, hare, pika, rabbit

Reptile/Amphibian: alligator, chameleon, crocodile, frog, gecko, iguana, lizard, newt, salamander, snake, toad, tortoise, turtle

Rodent: beaver, chinchilla, chipmunk, gerbil, gopher, groundhog, guinea pig, hamster, hedgehog, marmot, mole, musk rat, porcupine, rat, squirrel, woodchuck

Weasel: badger, ferret, marten, mink, mongoose, otter, polecat, skunk